

The international journal of science / 20 February 2020

nature

TWISTED NERVES

Solid tumours manipulate neurons
to promote cancer growth

Smart thinking

Time to set measurable
targets to protect
ecosystems

Sweet success

Light-driven synthesis
makes rare sugars from
common counterparts

Balancing act

How RNA silences
one X chromosome
in female mammals

Re-use of the X chromosome
in female mammals
nature.com

New biodiversity targets cannot afford to fail

Global goals to protect natural systems will be revised this year. China must help to ensure the new targets are measurable and meaningful.

Most measures of biodiversity suggest that things are going badly wrong. Some one million plant and animal species face extinction, according to the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES). And French President Emmanuel Macron last week called the battle for biodiversity and climate change “the fight of the century”.

A decade ago, countries united to create a 10-year plan, sub-divided into 20 targets, for protecting and conserving natural systems. The plan, also known as the Aichi Biodiversity Targets, expires at the end of this year – and most of the targets will not have been reached. Between 24 and 29 February, representatives of the international community will meet in Rome to discuss a new plan. A lot is at stake, and it's vital that the world unites behind the effort.

The meeting will consider a draft of an updated set of global goals, which must be agreed by the summer. Then, in October, world leaders will gather in Kunming, China, for the Conference of the Parties to the United Nations Convention on Biological Diversity. China will be in the chair, the first time it will lead on a conference of the parties to one of the ‘big two’ global environmental conventions (see News page 345).

These discussions are as important to biodiversity as the Paris agreement was to climate, and are likely to be similarly fraught. Conservation groups back more stringent and more measurable targets. European countries sit somewhere between the United States – which has long refused to sign the biodiversity convention – and developing countries, which will be looking to China to fight their corner. But China's efforts to build consensus have been set back by the coronavirus, which has seen parts of the country closed down.

To be fair, not every biodiversity policy has failed. Among the hard-won achievements is the 2014 Nagoya Protocol, an agreement stating that the benefits of genetic resources must be shared equitably among all of those – including Indigenous communities – who have contributed to their development. This can take time, and the World Health Organization has been discussing how to reduce potential delays when genetic information needs to be shared in public-health emergencies. But the protocol's existence is a win for multilateral science and environmental diplomacy.

By contrast, there's been no clear progress on the

 **Biodiversity is rarely allowed to stop or delay a new airport runway or power plant.”**

headline ambition to slow and eventually reverse the loss of biological diversity around the world.

The Aichi targets failed, in part, because their format makes progress hard to measure. Ahead of this year's talks, a group of researchers led by Elizabeth Green at the Centre for Conservation Science in Sandy, UK, scanned the literature for mentions of the Aichi targets since 2010. The team then invited an expert group to score the targets on a scale of one to ten. All of the targets scored highly for being comprehensive, but most scored relatively poorly on being measurable and realistic (E.J. Green *et al. Conserv. Biol.* 33, 1360–1369; 2019).

Take the first target, intended to ensure that “people are aware of the values of biodiversity and the steps they can take to conserve and use it sustainably”. It's clear this aims to raise public awareness of and engagement with biodiversity, but it's not clear when success has been achieved.

Those drawing up a new generation of biodiversity goals and targets understand this. The text of a new draft released last month contains spaces in square brackets, ready to be filled in when more-quantitative measures are agreed. Such measures include ensuring strict protection for important ecosystems and finding nature-based solutions that increase resilience to natural disasters (see Comment page 360).

Ambition versus achievement

The Aichi targets didn't fail solely because they weren't measurable. They also failed because countries did not need to report what they were doing to achieve them.

The biodiversity convention's member states have to publish biodiversity action plans – but these are often statements of a country's ambitions, rather than records of its achievements. For the next set of goals this has to change, and fortunately there seems to be a way forward. This is the UN System of Environmental Economic Accounting (SEEA), a mechanism for reporting environmental data, and it needs to become the global standard for environmental reporting.

SEEA was adopted in 2012 to encourage countries' national statistical offices to take responsibility for collecting and reporting environmental data. Asking statistics offices to do this was a stroke of genius. These offices are already responsible for reporting national economic data to the UN. They work to the best available standards and strict deadlines – and they get the job done. Charging them with reporting environmental data ensured that these data would be treated in the same way.

What began as a trickle of countries following the system has surged to more than 80 states sending updates to the UN on a multitude of environmental indicators, from the state of their forests to the state of their fisheries. Developing countries will need to be supported to get up to speed and contribute their own ideas. But the die is cast.

As is sometimes the case with the UN, a lack of joined-up thinking allowed SEEA to emerge independently of other indicators, such as the Aichi targets and the Sustainable Development Goals (SDGs). Now, moves are under way towards some harmonization. Last July, the UN published a global indicator review (go.nature.com/2ssazbc) in which

researchers confirmed that countries could use SEEA to report 34 of the 147 Aichi target indicators and 21 of the 230 SDG target indicators. This is an important start, but also indicates how much needs to be done before more goals and targets can be reported using the SEEA framework – an opportunity which researchers must not pass up.

Measuring and reporting numerical targets, although vital, is not the whole story. If the world is to understand why the Aichi targets failed – and improve on them – it must assess the broader obstacles.

One is the historical tension between development and the environment – and the expectation of poorer countries that they should be able to develop, just as richer countries did. There is also a perception that new environmental standards will hold them back. No one can contest their case for developing, but, considering the state of the planet, their concerns need to be met through greener development. They need support to provide their citizens with basic amenities – such as clean water, nutrition and power – in a way that is sustainable and protects future generations. This means making significant changes to how economic decisions are made.

No contest

Usually, in any contest between industrial growth and the preservation of species and ecosystems, growth comes out on top. Biodiversity is rarely allowed to stop or delay a new airport runway or power plant. If a wetland needs to be concreted over to make way for a housing development, in many countries it has little chance of being protected, even though losing the wetland means sacrificing the services it provides to people – such as wildlife habitats and flood defences. These services are rarely quantified.

Fortunately, researchers and policymakers globally are taking a stronger interest in valuing biodiversity's contribution to economies and to societies. IPBES is deep in a project that will advise countries on the many ways to value biodiversity; a report is due to be presented next year. And last year, the UK Treasury launched its own independent review, chaired by the economist Partha Dasgupta of the University of Cambridge, that is due to report in time for the biodiversity conference in China.

We know that working in an economic and financial system that places little value on the natural world will make it difficult to meet goals in biodiversity and sustainable development. That's why it is prudent to tackle smaller aspects of the system – at least for now. At the same time, it's imperative that the new biodiversity goals find synergies and avoid conflicts with the Paris climate agreement and the SDGs, neither of which existed a decade ago.

The road to the Kunming convention will be long and complicated. This is inevitable, both because life on Earth is itself beautifully complex, with so many global systems influencing biodiversity, and because the outcomes matter. Humanity's future depends on our ability to protect the planet. Greater awareness of threats to the natural world – perhaps an intangible impact of the Aichi targets – has created a moment ripe for action. The challenge will be to keep the devil in the detail from derailing the process itself.

“
The REF's
critics need
to be careful
what they
wish for.”

The final countdown

The United Kingdom's Research Excellence Framework might turn out to be the last.

It was the day most UK academics were dreading. On Monday 17 February, funding agencies fired the starting gun on the next Research Excellence Framework (REF 2021), the United Kingdom's system for evaluating research quality.

Universities have until 27 November to submit their researchers' outputs to the REF. These will then be graded by review panels on a scale of 1 to 4 – the highest score meaning that the work is deemed “world leading” in its originality, significance and rigour.

A lot is riding on the outcome because funders use the results to allocate around £2 billion (US\$2.6 billion) in annual research funding to university departments. Most institutions will want to see their academics graded in the top two bands, because lower-performing departments are unlikely to get much money at all.

The exercise is valuable in providing public accountability for research spending while protecting universities' financial autonomy. But many researchers and research managers are wondering whether REF 2021 could be the last.

Many would not mourn the REF's demise. By coincidence, from 20 February thousands of UK academics will be on strike for 14 days, calling for better pay and more-secure pensions. The constant monitoring of performance that comes with research evaluation is also mentioned by academics as a source of stress and anxiety.

The REF is also not cheap to administer – the 2014 exercise cost around £246 million. And as with most indices, the REF's overlords keep having to make changes to prevent it from being gamed. In the past, departments were able to achieve high scores by submitting outputs from a fraction of their best-performing staff – something that is no longer allowed.

Universities that obtain the most REF-based funding are concentrated in London and southeast England, and this has fuelled arguments that the metric's funding formula helps to reinforce the UK's regional imbalance. That alone could be an argument for radical reform from a government looking to level up funding to other parts of the United Kingdom.

That said, the REF's critics need to be careful what they wish for, because the framework protects money that universities rely on to pay salaries and to keep the lights on. The government of Prime Minister Boris Johnson is also looking to cut funding from publicly funded bodies that have operated largely autonomously from the state – including the national broadcaster, the BBC. Moreover, proposals for research funding reforms are widely expected this year.

A bonfire of the REF might well appeal to many, but not if the outcome leads to cuts, or reduced autonomy for institutions. There could be a wiser option: adjust the REF's funding formula so that money for the best work is distributed more fairly across the United Kingdom.

World view



By Maggie
Ryan Sandford

You can't fight feelings with facts: start with a chat

I donned a sandwich board inviting questions on evolution and learnt three crucial lessons about public engagement on divisive issues, writes Maggie Ryan Sandford.

I went to the Minnesota State Fair last year wearing a sandwich board. It said, “Ask me anything about evolution.” Proponents of evolution assumed I was a religious zealot. Creationists assumed I was there to mock their beliefs. The biggest challenge in fighting misinformation? Just getting a conversation started.

This public-engagement stunt taught me a crucial lesson: the key to effective science communication isn't the science. It's communication.

Attendees had come to show off prize livestock, eat corn dogs and ride the Ferris wheel, not get angry about someone who disagrees with them about the origin of life on Earth. Most folks wouldn't stop to talk unless I passed what I came to recognize as ‘the first test’. Some would call out, without slowing: “Do you believe in evolution?” Others, “Do you believe in God?”

Part of me died each time I answered with a profoundly un-nuanced “Yes!” But, as a science communicator and former education researcher, I knew that, in matters of deep personal belief, facts matter less than feelings. The need to identify whom you're dealing with is a natural human instinct. Answering was the only way to unlock the rest of the conversation. So I simply let people know I was a big fan of the globe and everything on it, and that I'd written a book about animals that I hoped people would find inviting.

In the cow barn, a man with his pre-teen daughter smiled at me, then avoided my gaze. “We don't believe in evolution, so ...” “Okay!” I called back. “Did you want to talk about cows?” I pointed out how humans can build muscle by eating cow protein, because of our shared ancestry. We know that ‘relevancy’ is crucial to public understanding – people need takeaways that relate to their everyday lives. “Well, that's awesome,” he said, “because I do love a good steak!” Before he and his daughter walked away, we exchanged thumbs ups.

Lesson 1: Don't argue with beliefs. People tend to incorporate facts that align with their belief systems.

No problem. I just had to find topics that made sense to all of us – pro-and anti-evolution alike. Dogs or livestock breeding, for example. Half the folks within a 30-metre radius were there to showcase their carefully bred cows, horses and chickens. Open-faced and genuine, I invited them to school me on the areas of their expertise. Which, it turns out, is evolution.

Lay people are more likely to trust and engage with science when they learn that researchers are human beings, fallible and conflicted.

Maggie Ryan Sandford is a science communicator and author of the 2019 book *Consider the Platypus: Evolution through Biology's Most Baffling Beasts*. e-mail: m.r.sandford@gmail.com

Lesson 2: Listen. The most challenging group of the day consisted of two men and a woman in their late twenties. The men were just looking for a fight. Telling me why I was wrong was, I supposed, a way of asking me about evolution. I asked them to elaborate, to tell me why it was that they found evolution hard to swallow. This led to their female companion insisting: “She listened to you. Now you listen to her.” In the end, one man explained my points to the other. “She's saying evolution is mutations in our DNA,” he said, forcing his companion to let him finish. “I'm just saying, I get her side.”

Lesson 3: Learn what people really think. Almost everyone – secular and religious – had misconceptions about evolution. Advocates of evolution often hadn't learnt that evolution can now be tracked in genomes, not just fossils, and that humans are related to all living things, and that we didn't come from apes because we are apes (keep in mind, ‘ape’ is a word that humans made up).

But the misconceptions of religiously inclined folks often had greater personal significance. Listening to them, it became clear that they considered evolution an attack on all they held dear. Several asked me about a narrative they'd heard somewhere about how “life began when water was dripping on a rock”. Clearly, they were worried that such a narrative undercut the idea that humans were created in the image of God.

People from both groups often misinterpreted the term ‘survival of the fittest’, and were surprised to hear that evolution isn't a system of improvement, just a system of change. And that *On the Origin of Species* was not intended as an attack on faith. Even in old age, Darwin declared: “I have never been an atheist.”

Lay people are more likely to trust and engage with science when they learn that researchers are human beings, fallible and conflicted. Yet somehow it seems hard for many in the scientific community to show those qualities to others. A common concern is that, in the anti-evolution, anti-science debate, any whiff of disagreement or uncertainty spells doom for scientific arguments.

When I began this ‘experiment’, my hypothesis was that a willingness to show vulnerability – to show that we science folks are willing to listen and receive criticism – boosts credibility, not the opposite. I think my experience supports that. When feelings speak louder than facts, appealing to feelings can actually work in favour of science.

No matter where we think we fall in the evolution debate, all of us are human, and we evolved to read each other's facial expressions and tones of voice, to be together. Returning to our humble, apish roots is the only way to see anti-science sentiment go the way of the dodo.

News in brief

ANIMAL-RESEARCH DATA SHOW EFFECTS OF EU'S TOUGH REGULATIONS

Scientists in the European Union seem to be using fewer animals for research, according to statistics gathered by the European Commission. The figures come from the first report on the state of animal research in the bloc since the introduction of tougher regulations seven years ago.

The report – published on 6 February – reviews the impact of an animal-research directive, legislation that was designed to lessen the use of animals in research and minimize their suffering. The directive is widely considered to be one of the world's toughest on animal research.

According to the report, 9.39 million animals were used for scientific purposes in 2017 – the most recent year for which data have been collated – compared with 9.59 million in 2015. From 2015 to 2016, however, there was a slight increase, to 9.82 million. The report acknowledges that this prevents the confirmation of a clear decrease. But it adds that, when compared with figures from before the directive came into force, the numbers suggest

“a clear positive development”.

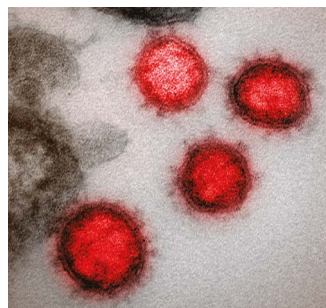
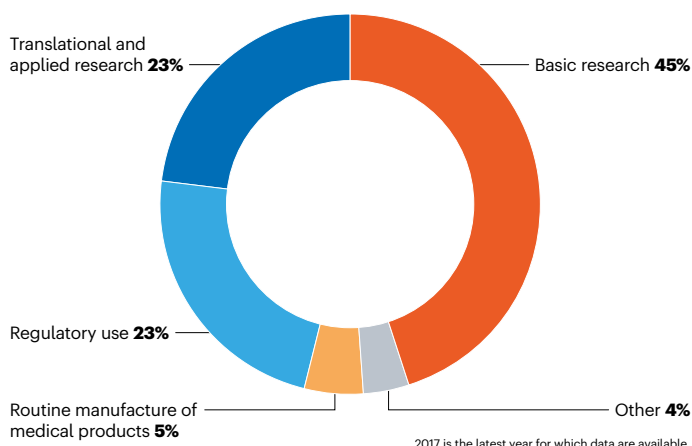
In 2017, more than two-thirds of instances of animal use were in basic or applied research (45% and 23%, respectively), and around one-quarter (23%) involved the testing of drugs and other chemicals to meet regulatory requirements. Other uses included the routine production of biological agents such as vaccines; teaching; and forensic investigations (see ‘Animals in science’).

The legislation sets out high standards for the housing and care of animals, and promotes methods that cause the least pain and use a minimal number of animals. It requires member states to submit detailed data, including the number and species of animals used in research, as well as the number of times each animal is used, and the purpose and severity of experimental procedures.

A spokesperson for the European Commission says that such detailed data “allow us to identify far more effectively where best to target resources to help reduce the number and suffering of animals”.

ANIMALS IN SCIENCE

In 2017, more than two-thirds of recorded instances of animal use in the European Union were in basic or applied research.



CORONAVIRUS NAME PROMPTS CONTROVERSY

The disease caused by the new coronavirus now has an official name: COVID-19.

The disease, and the virus, had been going by a number of monikers, including 2019-nCoV, since they emerged in China in December. The World Health Organization (WHO), which announced the new name on 11 February, said that it chose one that did not refer to a geographical location, an animal or a group of people, to avoid stigma.

On the same day, a group with the International Committee on Taxonomy of Viruses, which is responsible for naming the pathogens, designated the virus itself SARS-CoV-2 (A. E. Gorbalenya *et al.* Preprint on bioRxiv <http://doi.org/dmsh; 2020>). The group said that this term highlights the new virus's similarity to SARS-CoV, the coronavirus identified in 2003 that causes severe acute respiratory syndrome.

But the virus name caused consternation, particularly among Chinese virologists, who worry that it will confuse the public and impede efforts to control the pathogen's spread. Although the two viruses belong to the same species, the new coronavirus spreads faster than SARS-CoV but is less deadly, says Shibo Jiang, a virologist at Fudan University in Shanghai. The new coronavirus has infected more than 73,000 people.

INFLUENTIAL CLIMATE CHIEF DIES

Rajendra Pachauri, an Indian environmentalist and former head of the Intergovernmental Panel on Climate Change (IPCC) who had been accused of sexual harassment, died on 13 February. He was 79 and his death followed recent heart surgery, according to media reports.

From 2002 to 2015, Pachauri was chair of the IPCC – the international organization that produces scientific reports on the state of climate change and developed the Paris agreement to halt global warming. In 2007, during his tenure, the organization received the Nobel Peace Prize.

Born in 1940 in Uttarakhand state, Pachauri studied engineering and economics in India and the United States. He became director of the Energy and Resources Institute (TERI), a climate and energy-policy institute based in New Delhi, in 1981. He received several civilian honours from the Indian government.

But in 2015, he stepped down as chief of the IPCC and from TERI's leadership after a female colleague accused him of sexual harassment. Pachauri denied the accusations; a case was pending in a Delhi court at the time of his death.





Solar System's distant snowman comes into sharp focus

The distant Solar System object known as Arrokoth resembles a reddish snowman, data from NASA reveal.

NASA's New Horizons spacecraft – which travelled to the farthest reaches of the Solar System – took this image in January 2019, when it flew past Arrokoth. The object, which was previously known as 2014 MU₆₉, lies beyond Pluto in the frigid Kuiper belt, and is the most distant Solar System object imaged up close.

Scientists with the mission published new findings on the rocky object in *Science* on 13 February. They show that Arrokoth's two lobes are not quite as flat as they appeared previously, and that they probably merged gently in the early days of the Solar System, at least four billion years ago. Arrokoth, which is 36 kilometres long, is extremely red, probably because cosmic rays have blasted its surface to create red organic molecules. Unlike many objects in the outer Solar System, Arrokoth has no water frozen on its surface, although it does have methanol ice (see go.nature.com/37evuex for more).

It is probably typical of the Kuiper belt objects in similar orbits, says David Jewitt, an astronomer at the University of California, Los Angeles. But it would take another spacecraft visit to find out conclusively, he says. "We'll never know for sure until we look."



UK SCIENCE MINISTER OUSTED

A government reshuffle last week ousted the UK minister for universities and science, Chris Skidmore.

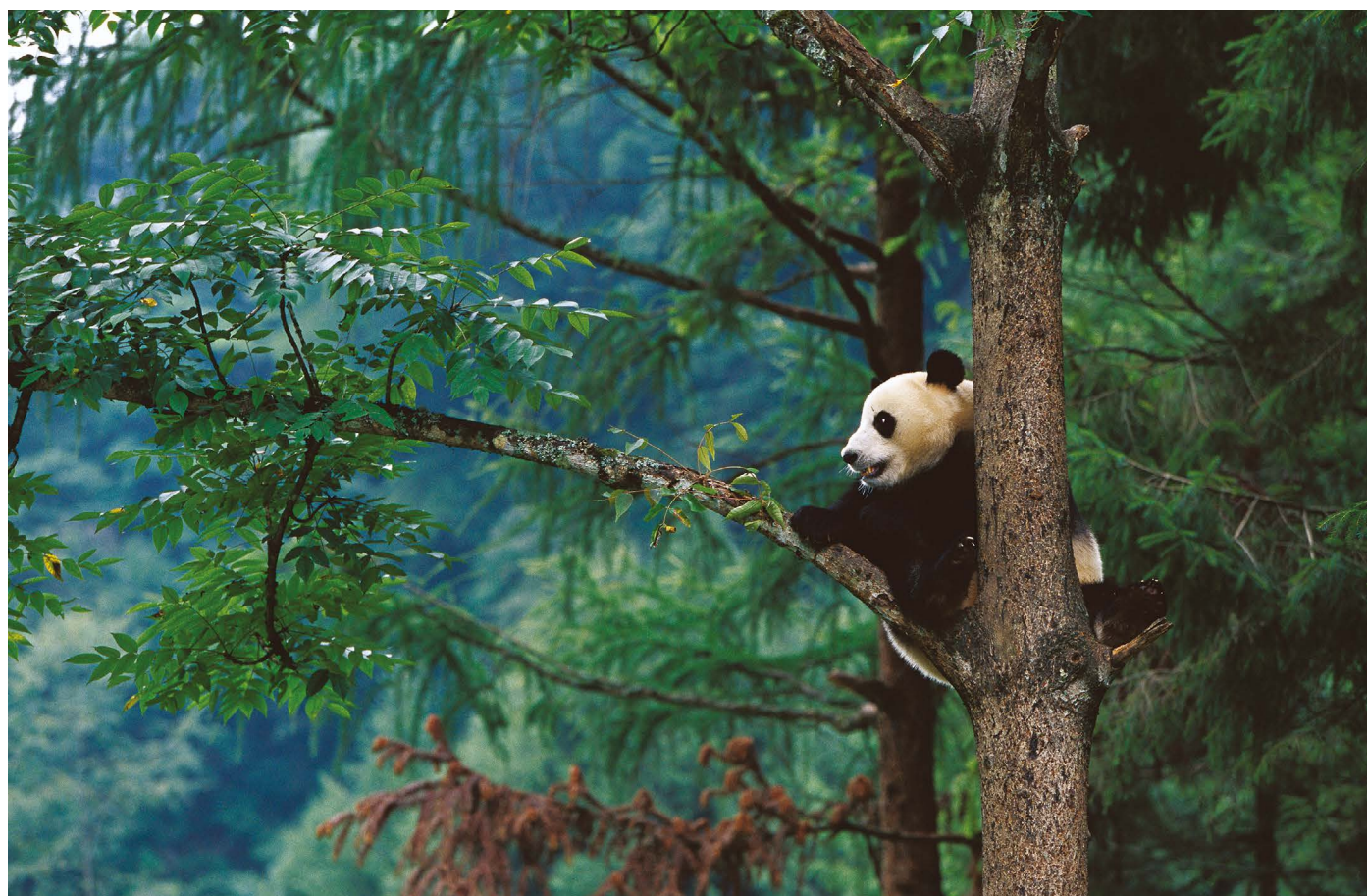
Skidmore had occupied the position – which has seen a revolving door of appointees and a series of resignations in recent years – for two periods since 2018, and was popular with academics.

As *Nature* went to press, it was not clear who would take on the ministerial briefs for universities and science. But preliminary appointments by Prime Minister Boris Johnson's Conservative government suggested that responsibility for the areas would be split between ministers for the first time since 2010.

Alok Sharma (pictured) was appointed secretary of the Department for Business, Energy and Industrial Strategy, which oversees science. The government also named Sharma president of COP26, the high-profile climate conference to be held in Glasgow in November, where nations will review their pledges to cut emissions.

The reshuffle comes as the United Kingdom heads into negotiations over its future relationship with the European Union. Crucial for UK scientists will be whether they can take part in Horizon Europe, the EU's next major research-funding programme, following the United Kingdom's exit from the bloc last month.

News in focus



CYRIL RUOSO/MINDEN PICTURES/NATL GEOGRAPHIC

The giant panda (*Ailuropoda melanoleuca*) was upgraded from endangered to vulnerable in 2016.

CHINA TAKES CENTRE STAGE IN MAJOR BIODIVERSITY PUSH

A United Nations summit could see China press for ambitious targets, and spotlights the country's own conservation efforts.

By Smriti Mallapaty

The world's species and natural ecosystems are in crisis. When nearly 200 countries gather next week to thrash out a major plan to stem the precipitous decline, China is expected to take a prominent role. The high-stakes negotiations will set the stage for a major biodiversity summit in October, which the country will host – marking the first time the nation will lead global talks on the environment.

That role, together with China's growing

global influence – including its vast Belt and Road Initiative to build international infrastructure – has put a spotlight on its impact on, and efforts to preserve, biodiversity.

“We are familiar with China being part of the problem of the global environmental emergency. For the sake of nature and the people living on this planet, there is a need to turn China into part of the solution,” says Li Shuo, a policy adviser at Greenpeace China in Beijing.

The gathering on 24–29 February was originally planned to take place in China, in the city of Kunming, but following the outbreak

of coronavirus in December, it has been moved to Rome.

The meeting is the second of three rounds of talks in which nations decide biodiversity targets that will form the basis of a legally binding global agreement. That will be signed at the 15th conference of the parties to the United Nations Convention on Biological Diversity (CBD) in Kunming in October, and will replace the current accord, signed in 2010.

The CBD global agreements are the main mechanism for holding signatories responsible for protecting biodiversity for the next

News in focus

ten years, says Basile van Havre, co-chair of the CBD. More than 190 countries have signed the treaty.

The stakes are particularly high this time, because countries have largely failed to meet the 2020 deadline for the current CBD goals, such as preventing species extinctions and ensuring that all fish stocks are harvested sustainably, says Li.

Ecosystems are vanishing rapidly, and close to one million plant and animal species face extinction. If this trend continues, it could have big consequences for people and food production, says van Havre. Countries have failed to meet the current goals partly because the targets were vague and difficult to implement, and progress hard to track, he says.

As host, China will take over presidency of the CBD and lead negotiations at the October summit. The country's diplomats will have the crucial job of nudging the world to agree to ambitious goals that can be measured, say researchers. At the February meeting, China is expected to take a prominent role in negotiations over amendments to a paper, known as the zero draft, that will form the basis of the accord to be signed in October.

China has made significant progress on environmental issues at home, which will give its diplomats some clout, says Li. Over the past decade, the central government has established thousands of nature reserves and parks, and it is drawing up ecological 'red lines' to restrict human and industrial activity over about one-quarter of the country. And in 2016, the conservation status of the giant panda (*Ailuropoda melanoleuca*) was upgraded from endangered to vulnerable. "Hosting this conference is one way of demonstrating that China is getting more serious on environmental

protection," says Li.

But China has also had high-profile failures. Last December, a group of scientists said the giant paddlefish (*Psephurus gladius*) that swam in the waters of the country's Yangtze River was extinct (H. Zhang *et al. Sci. Total Environ.* **710**, 136242; 2020). And no one has seen a Yangtze River dolphin (*Lipotes vexillifer*) since 2002; the species might also be extinct.

For China to truly lead when it comes to biodiversity, it will need to address its significant and rising ecological footprint outside its borders, says Aleksandar Rankovic, an environmental scientist at the Institute for Sustainable Development and International Relations in Paris.

"This conference is one way of demonstrating that China is getting more serious on environmental protection."

The country imports large volumes of environmentally damaging products such as palm oil, which contributes to deforestation in tropical countries including Indonesia. And although trends are hard to estimate, much of the world's illicit wildlife trade is widely understood to be driven by Chinese demand for traditional medicines and delicacies.

Li Zhang, a conservation biologist at Beijing Normal University, thinks that China's wildlife trade and consumption contributed to the current coronavirus outbreak. The virus is thought to have jumped to humans from animals, possibly pangolins, sold in a market in Wuhan, the epicentre of the outbreak. The national government has temporarily banned

wildlife markets. But Li says these measures are not enough to prevent the emergence of another infectious disease, and is calling for a permanent ban on the markets.

Infrastructure impacts

Conservationists are also worried about the environmental impact of China's Belt & Road Initiative (BRI), a vast infrastructure-building project to connect China to many parts of the world that has been likened to a modern Silk Road. A 2019 study found that many of the road and rail lines planned for southeast Asia pose a serious risk to biodiversity hotspots and could facilitate wildlife trafficking (A. C. Hughes *Conserv. Biol.* **33**, 883–894; 2019). Belt and Road projects, especially in developing countries, have not had enough oversight to ensure that ecologically fragile regions are protected, says study author Alice Hughes, a conservation biologist at the Chinese Academy of Sciences Xishuangbanna Tropical Botanical Garden.

The BRI is not aligned with the biodiversity-conservation approach that researchers hope China will champion as part of the Kunming agreement, says Simon Zadek, the London-based principal of Project Catalyst, an initiative of the United Nations Development Programme to spur action towards achieving the UN Sustainable Development Goals.

Many countries are using the Kunming conference in October to highlight the environmental impact of the BRI, says Li at Greenpeace China. "This will be a big thing to watch out for, whether there will be an effective policy response from China's side," he says.

Hughes says researchers are calling for the biodiversity treaty to address the international environmental impacts created by all countries – not just China – through activities such as foreign investment and trade. "International development projects are not being subjected to any level of scrutiny," says Hughes.

But van Havre says that the CBD doesn't have the scope to consider environmental damage from foreign-investment projects. It deals only with sovereign states and their responsibilities within their own borders, he says.

The negotiations in February will consider domestic mechanisms to protect the environment, and this year's zero draft is the first to include a specific mention of the need for countries receiving foreign investments to assess their environmental impact, says van Havre. For instance, if China wants to build a railway in another country, then that country needs to do such an assessment before allowing the project to go ahead, he says.

Li doesn't think China will address environmental concerns about the BRI seriously at this year's talks. The officials calling the shots on the BRI are not the ones handling the Kunming conference, he says. "I expect a lot of buzz on overseas footprint, but I do not expect a lot of concrete policy progress," he says.



The Yangtze River dolphin (*Lipotes vexillifer*) is feared to be extinct.

MARK CARWARDINE/NPL



Medics check on people with COVID-19 in Jinyintan Hospital in Wuhan, China.

SLEW OF TRIALS LAUNCH TO TEST CORONAVIRUS TREATMENTS IN CHINA

HIV drugs, stem cells and traditional Chinese medicines are vying to prove their worth.

By Amy Maxmen

China has more than 80 running or pending clinical trials on potential treatments for COVID-19, the illness caused by a coronavirus that has so far killed more than 1,800 people and infected more than 70,000 across the country, and for which there is currently no cure.

New drugs are listed beside thousand-year-old traditional therapies and existing treatments for other diseases in a public registry of China's clinical trials that is growing every day. But scientists caution that only carefully conducted trials will show which measures work.

Soumya Swaminathan, chief scientist at the World Health Organization (WHO), says that the agency is drawing up a plan for a clinical-trial protocol that researchers around the world could use, and working with scientists to help set standards for the trials in China, which include as many as 600 people each.

For example, a person's stages of recovery or decline should be measured in the same way, regardless of the treatment being tested, says Swaminathan. "We can hopefully bring some sort of structure into the whole thing."

The WHO's clinical-trial protocol will compare two or three therapies, including

an HIV-drug combination (lopinavir and ritonavir) and an experimental antiviral called remdesivir.

Researchers in China have begun testing these drugs in clinical trials, according to the Chinese Clinical Trial Registry, and there is already some evidence to suggest they have potential to fight the coronavirus. "Getting the clinical trials straight is a priority, since if we get information on what is working and not working, we can benefit patients now," Swaminathan says.

Animal results

The two HIV drugs block enzymes that viruses need to replicate. In animal studies, they have reduced levels of the coronaviruses that cause severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS)¹. Remdesivir, a nucleotide analogue made by the biotechnology company Gilead in Foster City, California, has also had some success against coronaviruses in animals². And in January, researchers reported that one person in the United States had survived a COVID-19 infection after being treated with remdesivir³.

In the first week of February, two placebo-controlled trials of remdesivir, set to include a total of 760 people with COVID-19,

began in China. Those studies should be completed by the end of April, and remdesivir could be approved by Chinese authorities as early as May, says Shibo Jiang, a virologist at Fudan University in Shanghai. "But the epidemic might be gone by then," he says.

Researchers in China have also launched a few trials that test chloroquine, a malaria drug that killed off the new coronavirus (recently named SARS-CoV-2) in cell culture⁴. And scientists are studying whether steroids diminish inflammation in people with severe COVID-19, or cause harm. "It will be interesting to see these results," says Yazdan Yazdanpanah, an epidemiologist with France's national health agency, INSERM, in Paris. Research clinicians around the world will need this information if the outbreak continues to spread, he adds.

Another study – a 300-person controlled trial – will test serum from COVID-19 survivors. The same basic idea – that the antibodies one person steadily builds up to fight a virus can help someone freshly infected to fight it off rapidly – has had modest success when used to treat other viruses in the past⁵.

Two stem-cell trials are also listed in China's registry. In one, a team at the First Affiliated Hospital of Zhejiang University will infuse 28 people with stem cells derived from menstrual blood, and compare results with those from people who did not receive the infusions. So far, there is minimal evidence indicating that stem cells clear coronavirus infections. Swaminathan says that the WHO cannot control what researchers do, but that the agency published guidance on the ethics of running trials amid outbreaks in 2016. And it will be posting a more accessible brief report on the issue soon.

About 15 trials listed in China's registry expect to enrol a total of more than 2,000 people in studies on a variety of traditional Chinese medicines. One of the largest assesses shuanghuanglian, a Chinese herbal medicine that contains extracts from the dried fruit *lianqiao* (*Forsythiae fructus*), which is purported to have been used to treat infections for more than 2,000 years. The trial has 400 participants, including a control group given standard care but not a placebo therapy.

The WHO is working with Chinese scientists to standardize the design of all the studies, including those on traditional medicines. That reflects a controversial move last year, in which the organization recognized traditional Chinese medicine in its compendium of diseases. Critics argued that the WHO's recognition amounted to endorsement, but Swaminathan disagrees. She says that the move helps to codify medical terminology so that herbal remedies can be evaluated with the rigour expected of pharmaceutical testing. "We want a scientific approach to testing traditional medicine," she says.

With many therapeutic possibilities and limited time, Jiang says the WHO should provide advice about which treatments to move forward, and which to ditch, as trials progress. And he hopes that research on better, broader therapies will be continued after the outbreak ends. “I worry this will be the same situation as during SARS,” he says,

“where the work starts, then stops.”

1. Guangdi, L. & De Clercq, E. *Nature Rev. Drug Discov.* <http://doi.org/10.1038/d41573-020-00016-0> (2020).
2. Sheahan, T. P. *et al. Nature Commun.* **11**, 222 (2020).
3. Holshue, M. L. *et al. N. Engl. J. Med.* <http://doi.org/10.1056/NEJMoa2001191> (2020).
4. Wang, M. *et al. Cell Res.* <https://doi.org/10.1038/s41422-020-0282-0> (2020).
5. Marano, G. *et al. Blood Transfus.* **14**, 152–157 (2016).

SCIENTISTS FEAR CORONAVIRUS SPREAD IN VULNERABLE NATIONS

Concerns are rising about the virus’s potential to circulate undetected in Africa and Asia.

By Smriti Mallapaty

Infections with the new coronavirus have now been detected in 25 countries outside China. But researchers warn that cases might be going undetected in some nations that are considered to be at high risk of an outbreak but are reporting fewer cases than expected, or none at all.

The possibility of unreported cases of the disease, known as COVID-19, is particularly concerning in countries with weaker health-care systems, such as some in southeast Asia and Africa, which could quickly be overwhelmed by a local outbreak, experts say. So far, only one case has been reported in Africa – in a person in Egypt – but some countries there, such as Nigeria, are at particular risk because of their strong business ties to China.

Researchers have been using flight data to create models of the possible spread of the virus around the world. One model identified 30 countries or regions at risk of importing the virus on the basis of the large number of flights from Wuhan, the outbreak’s epicentre, and from other cities in China with many travellers from Wuhan.

Thailand is the country most exposed, according to the study, which was published on 5 February and used flight data from February 2018 (S. Lai *et al.* Preprint at medRxiv <http://doi.org/dmr4;2020>). Thirty-five people with the infection have been reported there so far, of whom 23 had been in China. But study co-author Shengjie Lai, an epidemiologist at the University of Southampton, UK, says the model estimates that Thailand probably imported 207 cases in the 2 weeks before travel into and out of Wuhan was restricted in late January.

Indonesia has not reported a single case so far, and yet the country is a popular destination

for Chinese tourists. Lai says it might have imported as many as 29 cases. Several other countries, including Malaysia, Vietnam, Cambodia and Australia, have also reported fewer cases than the model predicts, he says.

Although it’s possible that there have truly been no cases in Indonesia, infected people might have recovered before they were detected, says epidemiologist Andrew Tatem, a co-author of the study also at the University of Southampton. Undetected cases might also be spreading under the radar, he says.

Despite the predictions, Amin Soebandrio, an infectious-disease scientist and chair of the Eijkman Institute for Molecular Biology in Jakarta, says Indonesia has the capacity to detect the virus in people if it arrives.

But some countries in southeast Asia have limited numbers of health-care workers,

hospital beds, support staff and ventilators, and would struggle to respond to a surge in cases of the virus, says Richard Coker, a retired physician based in Bangkok.

Tedros Adhanom Ghebreyesus, director-general of the World Health Organization (WHO), said the agency’s decision to declare the outbreak a global health emergency was mainly due to concerns that the virus could spread in countries with weaker health-care systems.

What about Africa?

For that reason, infectious-disease researchers are also worried about the virus spreading among people in Africa. A large number of Chinese labourers work in Africa, and their travel between China and Africa is a possible route for transmission, says Marc Lipsitch, an epidemiologist at the Harvard T.H. Chan School of Public Health in Boston, Massachusetts.

Another model found that Egypt, Algeria and South Africa are the countries in Africa that are most at risk of the virus spreading. The analysis, published on 7 February, examined flights to Africa from Chinese cities that had reported infections, but excluded cities in Hubei province, where Wuhan is located, because of the lockdown that has restricted travel from many cities there since late January (M. Gilbert *et al.* Preprint at medRxiv <http://doi.org/dmr5;2020>).

But these three countries also have the capacity to respond effectively to an outbreak, says Vittoria Colizza, who models infectious diseases at the Pierre Louis Institute of Epidemiology and Public Health in Paris and is a co-author of the Africa study.

Colizza is most concerned about seven African nations that have a moderate risk of importing the virus, but whose weak health-care systems, low economic status or unstable political situation make them highly vulnerable. These are Nigeria, Ethiopia, Sudan, Angola, Tanzania, Ghana and Kenya.

Until two weeks ago, many African nations did not have laboratories that could diagnose COVID-19, and samples had to be tested abroad. But the situation is changing rapidly, says Colizza. Africa has gone from having only two labs with the capacity to confirm the virus to having at least eight, according to the WHO.

Three of the newly added labs are in Nigeria, says Chikwe Ihekweazu, director-general of the Nigeria Centre for Disease Control in Abuja.

Ihekweazu says Nigeria’s size, the volume of travellers it receives and its vibrant economy already make it vulnerable to importing an infectious disease, and that the country’s strong business ties with China pose a further risk.

Nigeria has ramped up screening of travellers from China. Ihekweazu says the worst-case scenario for the country would be if an infected person goes undetected and begins to infect others. “That is really what keeps me up at night,” he says.



The coronavirus responsible for COVID-19.

POPULAR PREPRINT SITES FACE CLOSURE BECAUSE OF MONEY TROUBLES

Repositories such as INA-Rxiv boost regional science, but paying for them is proving difficult.

By Smriti Mallapaty

The rise of preprint repositories has helped scientists worldwide to share results and get feedback quickly. But several platforms that serve researchers in emerging economies are struggling to raise money to stay afloat. One, which hosts research from Indonesia, has decided to close because of this funding shortfall.

INA-Rxiv, which was set up in 2017, was one of the first repositories to host studies from a particular region. Previous platforms served specific disciplines, such as arXiv for physical-sciences research. Other region-specific repositories followed, including ArabiXiv, AfricArxiv and IndiaRxiv. These repositories increase exposure for research from the regions, and facilitate collaborations, say their managers.

The servers are run by local volunteers but hosted online by the non-profit Center for Open Science (COS) in Charlottesville, Virginia. The centre's platform hosts 26 repositories, including some that are discipline-specific. In 2018, the COS told repository managers that it would be charging them maintenance fees from 2020. The charges start at about US\$1,000 a year, and increase as repositories' annual submissions grow.

The costs can be significant, particularly for repositories in emerging economies. Dasapta Erwin Irawan, a hydrogeologist at the Bandung Institute of Technology who helped to set up INA-Rxiv, says the repository received more than 6,000 submissions between July 2018 and June 2019, so the fees will come to about \$25,000 per year, which he cannot afford. After unsuccessfully trying to raise money

from the Indonesian government, he has decided to wind down the service and close it, although he has not yet set an end date.

Juneman Abraham, a social psychologist at Bina Nusantara University in Jakarta, says he will lose an important source of information on the latest Indonesian research when the repository closes.

Long-term survival

The COS decided to introduce fees so that it could sustain its hosting service in the long term; running it will cost about \$230,000 in 2020, says Brian Nosek, the centre's executive director. It used to rely on grants from private foundations, but they are no longer enough. Now the operating costs will be covered by a mix of grants and user contributions, he says.

About half of the repositories have committed to paying the fees so far; some are managed by organizations that have access to grants, whereas others have partnered with libraries for funding, says Nosek. The centre is committed to helping repositories, including INA-Rxiv, find funding, he says.

But Nosek acknowledges that repositories run in emerging economies are most likely to struggle to raise funds. The centre will be flexible about when groups pay this year's fees, but if no money is received, it will freeze services so that they can't accept submissions, he says.



**The week's best science,
from the world's leading
science journal.**

NATURE.COM/NATURE/PODCAST

nature

THE NEXT CHAPTER FOR AFRICAN GENOMICS

Nigeria is poised to become a hub for genetics research, but a few stubborn challenges block the way.

By Amy Maxmen





In the affluent, beach-side neighbourhoods of Lagos, finance and technology entrepreneurs mingle with investors at art openings and chic restaurants. Now biotech is entering the scene. Thirty-four-year-old Abasi Ene-Obong has been traversing the globe for the past six months, trying to draw investors and collaborators into a venture called 54Gene. Named to reflect the 54 countries in Africa, the genetics company aims to build the continent's largest biobank, with backing from Silicon Valley venture firms such as Y Combinator and Fifty Years. The first step in that effort is a study, launched earlier this month, to sequence and analyse the genomes of 100,000 Nigerians.

At a trendy African fusion restaurant, Ene-Obong is explaining how the company can bring precision medicine to Nigeria, and generate a profit at the same time. He talks about some new investors and partners that he's not able to name publicly, then pulls out his phone to show pictures of a property he just purchased to expand the company's lab space.

"My big-picture vision is that we can be a reason that new drugs are discovered," Ene-Obong says. "I don't want science for the sake of science, I want to do science to solve problems."

It's too soon to say whether he will succeed. But his ambitions would have been unthinkable a decade ago, when most universities and hospitals in Nigeria lacked even the most basic tools for modern genetics research. Ene-Obong, the chief executive of 54Gene, is riding a wave of interest and investment in African genomics that is coursing through Nigeria. In a rural town in the western part of the country, a microbiologist is constructing a US\$3.9-million genomics centre. And in the capital city of Abuja, researchers are revamping the National Reference Laboratory to analyse DNA from 200,000 blood samples stored in their new biobank. Studying everything from diabetes to cholera, these endeavours are designed to build the country's capabilities so that genetics results from Africa – the publications, patents, jobs and any resulting therapies – flow back to the continent.

The rest of the world is interested, too. Africa contains much more genetic diversity than any other continent because humans originated there. This diversity can provide insights into human evolution and common diseases. Yet fewer than 2% of the genomes that have been analysed come from Africans. A dearth of molecular-biology research on the continent also means that people of African descent might not benefit from drugs tailored to unique genetic variations. Infectious-disease surveillance also falls short,

meaning that dangerous pathogens could evade detection until an outbreak is too big to contain easily.

But Nigeria's genetics revolution could just as soon sputter as soar. Although the country is Africa's largest economy, its research budget languishes at 0.2% of gross domestic product (GDP). Biologists therefore need to rely on private investment or on funding from outside Africa. This threatens continuity: one of the largest US grants to Nigerian geneticists, through a project known as H3Africa, is set to expire in two years. There are other challenges. Human research in Africa requires copious communication and unique ethical consideration given the vast economic disparities and history of exploitation on the continent. And a lack of reliable electricity in Nigeria hobbles research that relies on sub-zero freezers, sensitive equipment and computing power.

Yet with a hustle that Nigerians are famous for, scientists are pushing ahead. Ene-Obong hopes to pursue research through partnerships with pharmaceutical companies, and other geneticists are competing for international grants and collaborations, or looking to charge for biotech services that are usually provided by labs outside Africa. Last November, Nnaemeka Ndodo, chief molecular bioengineer at the National Reference Laboratory, launched the Nigerian Society of Human Genetics in the hope of bringing scientists together. "When I look at the horizon it looks great – but in Nigeria you can never be sure," he says.

Building the foundation

Around 15 years ago, Nigerian geneticist Charles Rotimi was feeling dismayed. He was enjoying academic success, but would have preferred to do so in his home country. He had left Africa to do cutting-edge research, and he was not alone.

Many Nigerian academics move abroad. According to the Migration Policy Institute in Washington DC, 29% of Nigerians aged 25 or older in the United States hold a master's or a doctoral degree, compared with 11% of the general US population.

After Rotimi joined the US National Institutes of Health (NIH) in Bethesda, Maryland, in 2008, he hatched a plan with director Francis Collins to drive genetics research in Africa. Rotimi wasn't interested in one-off grants, but rather in building a foundation on which science could thrive. "The major thing to me was to create jobs so that people could do the work locally," he says. In 2010, the NIH and Wellcome, a biomedical charity in London, announced the H3Africa, or Human Heredity and Health in Africa, project. It's become a \$150-million, 10-year initiative that supports institutes in 12 African countries. The proof of its success will be not in the number of papers published, but rather in the number of African

54Gene aims to create Africa's largest biobank.



54Gene chief executive Abasi Ene-Obong is preparing to make Nigeria a genetics powerhouse.

investigators able to charge ahead after the grant ends in 2022.

For that to happen, H3Africa researchers realized they needed to revise research regulations and procedures for gaining the public's trust. So rather than just collecting blood and leaving – the approach disparagingly referred to as helicopter research – many investigators on the team have devoted time to adapting studies for the African context.

For example, when Mayowa Owolabi, a neurologist at the University of Ibadan, Nigeria, was recruiting healthy controls for his H3Africa study on the genetics of stroke, his team discovered that many people had alarmingly high blood pressure and didn't know it. Nigeria has one of the world's highest stroke rates, and Owolabi realized that communities needed medical information and basic care more urgently than genetics. So he extended his study to include education on exercise, smoking and diet. And, on finding that many people had never heard of genetics, the team attempted to explain the concept.

This is a continuing process. One morning last November – seven years into the project – a community leader in Ibadan visited Owolabi's private clinic. He said tensions had mounted because people who had participated in the study wanted to know the results of their genetic tests. Owolabi replied that they were still searching for genetic markers that would reveal a person's risk of stroke, and that it might be many years until any were found. "But it's a heart-warming question," he says, "because if the people demand a test, it means the study is the right thing to do."

Discovering the genetic underpinnings of stroke is also complicated by the fact that it, like many non-communicable disorders, is caused by a blend of biological and environmental factors. Owolabi flips through a blue booklet of questions answered by 9,000 participants so far. It asks about everything from family medical history to level of education. Insights are buried in the answers, even without DNA data: the team found, for instance, that young Nigerians and Ghanaians who eat green leafy vegetables every day have fewer strokes¹. And that's just the beginning. "You see the amount of data we have accrued," he says. "I don't think we have used even 3% of it, so we need to get more funding to keep the work going."

Owolabi's team is now applying for new grants from the NIH, Wellcome and other international donors to sustain the work after the H3Africa grant ends. And to make themselves more appealing to collaborators and donors, they're increasing the amount of work they can do in Ibadan. Until last year, most of the genetic analyses were conducted at the University of Alabama in Tuscaloosa. But last June, the University of Ibadan installed a computer cluster to serve the project, and three young bioinformaticians are now crunching the data. "The big-data business is happening now," says Adigun Taiwo Olufisayo, a doctoral student concentrating on bioinformatics. But he also admits that funds are tight.

Last year, other graduate students on the team began to extract DNA from samples so that they can scour it for genetic variants linked to strokes. In a room the size of a

cupboard, a technician labels tubes beside a freezer. Coker Motunrayo, a doctoral student studying memory loss after strokes, sits on the counter-top because there's not enough space for a chair. She insists that the H3Africa project is a success, even though their genetics work has just started. "Compare this to where we were five years ago, and you'd be stunned," she says.

On the cusp

Perhaps the most advanced genomics facility in West Africa right now is located in Ede in southwestern Nigeria. At Redeemer's University, a private institution founded by a Nigerian megachurch, microbiologist Christian Happi is building an empire. Construction teams are busy creating a \$3.9-million home for the African Centre of Excellence for Genomics of Infectious Diseases.

Happi strides across a veranda, and into a series of rooms that will become a high-level biosafety laboratory suitable for working on Ebola and other dangerous pathogens. Another small room nearby will house a NovaSeq 6000 machine made by Illumina in San Diego, California, a multimillion-dollar piece of equipment that can sequence an entire human genome in less than 12 hours. It's the first of that model on the continent, says Happi, and it positions his centre, and Africa, "to become a player in the field of precision medicine". Then he announces that Herman Miller furniture is on the way. If it's good enough for his collaborators at the Broad Institute of MIT and Harvard in Cambridge, Massachusetts, he adds, it is good enough for his team.

Happi plans to move his lab into the facility in a few months. But the team is already doing advanced work on emerging outbreaks. At a small desk, one of Happi's graduate students, Judith Oguzie, stares at an interactive pie chart on her laptop. The chart displays all of the genetic sequences recovered from a blood sample shipped to the lab from a hospital as part of a countrywide effort to learn which microbes are infecting people with fevers. Typically, doctors test the patients for the disease they think is most likely, such as malaria, but this means other infections can be missed. For example, the sequences Oguzie is looking at belong to the *Plasmodium* parasites that cause malaria, the virus that causes the deadly Lassa fever, and human papillomavirus.

Oguzie says that a few years ago, she was processing samples from a hospital in which people were dying because their fevers had confounded diagnosis. With the help of next-generation sequencing, she found that they were infected with the virus that causes yellow fever. She showed Happi the results, and he reported the news to the Nigeria Centre for Disease Control (NCDC), which rapidly launched a vaccination campaign.

This was exactly what Oguzie had wanted

out of science. “I’m happy when I solve problems that have to do with life,” she says. She worked hard throughout university in Borno, even after the terrorist organization Boko Haram started attacking the northern state. She heard bomb blasts during lectures and knew people who were shot.

Nevertheless, Oguzie finished her degree in 2011. She had a son a few years later and wanted to stay with her family in Nigeria, but she struggled to find a graduate institution that would allow her to excel in genetics. She had already begun searching for scholarships at universities in the United Kingdom, Australia and the United States when she found out about Happi’s lab.

Happi had been persuaded to return to Nigeria from Harvard School of Public Health in Boston. The vice-chancellor of Redeemer’s at the time was an influential virologist named Oyewale Tomori. He offered Happi a lucrative start-up package to build an environment similar to the one he had become used to at Harvard.

Soon after he joined the university, Happi won H3Africa grants totalling \$6.8 million that have led to some impressive projects. For example, he and his collaborators mapped the spread of infections in the country’s largest outbreak of Lassa virus². He also won World Bank funding for an African centre of genomics. The grant is paid out incrementally on the basis of milestones such as training graduate



BLOOD IS A RESOURCE, WHETHER IT’S INSIDE HUMANS OR OUTSIDE.”

students or researchers from another African country. So far, his centre has earned more than \$9 million.

He says the money means that he can offer experienced researchers salaries that stop them from leaving Nigeria and keep his lab up to date with the fast-moving field. Happi invites a rotating cast of top infectious-disease scientists from the United States to collaborate with his team in Ede. “I want to build a place where we can work together,” he says, “not a place from where things are taken away.”

But in an office beside Happi’s, geneticist Onikepe Folarin says she has no time to conduct research because she’s constantly writing grant proposals, and reporting back to donors on various milestones. To lessen their reliance on grants, she and Happi plan to start selling genomics services.

At the moment, African researchers pay a lot to ship samples and reagents to and from China and the United States, and these items often get held up at ports. But with his

sequencing equipment and machines to produce important reagents such as primers, Happi hopes to provide a commercial service to other researchers on the continent – and use the money to fund his research.

Disruptors

As the son of a plant geneticist, 54Gene head Ene-Onong developed a certain angst about the fits and starts of international grants. So after earning a PhD in genetics, he studied business with the aim of driving research sustainably. One idea he has for 54Gene is to charge drug-development firms for access to the genetic data in the company’s biobank. This model has proved successful elsewhere. For example, last year, the UK Biobank received \$120 million from 4 pharmaceutical giants for access to information on 125,000 people.

54Gene won’t say how it is financing its study to analyse 100,000 genomes from Nigerians, but it has gained the backing of physicians from 17 hospitals across the country, who will send blood samples from consenting people with chronic diseases such as cancer, diabetes and Alzheimer’s disease.

But as the first for-profit genetics endeavour in Nigeria, 54Gene must navigate uncharted ethical territory. People could feel cheated if they donate samples to research and then learn that the company turned a profit while they struggle to afford health care. Concerns about being taken advantage of loom large



Onikepe Folarin and Christian Happi stand in front of a soon-to-be completed genomics centre for studying infectious disease in Nigeria.



A perennial concern, Nigeria's underpowered infrastructure frustrates technology firms.

in Nigeria – and in Africa more generally – because of a history of the continent being exploited for everything from slavery to diamonds. As Anthony Ahumibe, the senior laboratory adviser at the NCDC, says: “Blood is a resource, whether it’s inside humans or outside.”

The concerns are well founded. Last year, for example, the Sanger Institute in Hinxton, UK, came under fire for licensing a gene chip based on African genome data to US biotechnology company Thermo Fisher, which was planning to manufacture the chip for a profit. This infuriated both the African researchers who had collaborated with the British team and the Ugandan study participants, who had not consented to the deal.

Seeing the potential for disaster, Aminu Yakubu, a bioethicist who helped revise Nigeria’s regulations at the start of the H3Africa projects, offered to join 54Gene last year to help the company come up with solutions. “I understand why people will be sceptical, so we will be as transparent as possible, and sensitive to concerns about exploitation,” he says. He and Ene-Obong are devising ways to give back to the public even before genetic discoveries are made. For example, they might donate dialysis machines to participating hospitals that lack them. “We are not just doing this to make money,” says Ene-Obong. “As a private company, we need money to operate, but my goal is to study African genetics and translate the insights into products that help people.”

The barriers

Unlike their younger colleagues, some established Nigerian researchers hesitate to celebrate the country’s inarguable growth in genomics because they see obstacles in the

path ahead. One of the biggest challenges is the lack of national funding. In 2016, it seemed that Nigeria’s government was realizing the importance of research when it approved a measure to commit 1% of its GDP to science and technology. That would have amounted to \$3.8 billion last year, but the money never materialized, and the research budget remains at about \$750 million annually – the total across all fields.

Tomori compares this situation with that in China – another middle-income country. A decade ago, China’s government plied the field of genetics with incentives such as tax exemptions and housing for scientists, and it



**IF WE SIT IN OUR LABS
DOING THE SAME
THINGS, THE SITUATION
WILL NOT IMPROVE.”**

put 2% of its GDP into research. Those investments have paid off; in 2018, China surpassed Europe in biotech investment.

And because the Nigerian government does not fund much science, it has limited power to set research agendas. That could stunt genetics projects because the most powerful studies stem from long-term national initiatives, such as the UK Biobank and the China Kadoorie Biobank, says Prabhat Jha, an epidemiologist at the University of Toronto in Canada. Nigeria does have a few large biobanks, generally attached to specific research projects – and 54Gene’s would add to that, but Jha warns that

it’s often difficult to cobble together samples from disparate studies because the data were collected with different aims. Creating a unified genomics initiative should be a priority, he says. “If there were good prospective studies under way in Africa,” he adds, “we could really start to understand the key determinants of diseases and deaths there.”

Even more basic problems stand in the way of success, not least the lack of a reliable electrical grid. “Until the government puts in basic infrastructure, we cannot move forward,” says Tomori. In the meantime, institutes and companies are spending a huge portion of their budgets on back-up generators, diesel fuel and solar panels. According to a report released last year by the International Monetary Fund, Nigeria’s inadequate electricity supply costs the country about \$29 billion per year³. And in a survey by the Center for Global Development, Nigeria’s booming tech sector named electricity as its number-one constraint⁴.

To change the status quo, Tomori says, his Nigerian colleagues must persuade their leaders and the public that investments in science matter. “If we sit in our labs doing the same things, the situation will not improve,” Tomori says. “We need to get out of our test tubes and talk about these issues.”

But the director of genomics research at the Nigerian Ministry of Science and Technology in Abuja, Oyekanmi Nash, argues that government funding will flow more freely once science starts to deliver tangible benefits. He credits H3Africa with triggering the first steps. Now, he says, it’s up to researchers to build on the effort and show how their science helps. Nash joined 54Gene’s initiative to sequence 100,000 genomes because of the start-up’s promise to translate genetics results into medicines. “Once we become strong enough,” he says, “the government will listen.”

It’s a tough bet to make, especially given that Nigeria’s post-recession economy remains sluggish. But the country’s younger geneticists don’t really have an option outside of optimism. “It’s not been easy,” says Ndodo. “Most of us have worked until the middle of the night, taken out loans to get training outside [Nigeria], and then come back to change the system.” But, he says, scientists are on firmer ground than their predecessors. And they’re driven. “No one else will tell our story,” Ndodo says. “No one else will do research that targets our own interests.”

Amy Maxmen writes for *Nature* from Oakland, California.

1. Sarfo, F. S. et al. *Stroke* **49**, 1116–1122 (2018).
2. Siddle, K. J. et al. *N. Engl. J. Med.* **379**, 1745–1753 (2018).
3. International Monetary Fund. *Nigeria: Staff Report for the 2019 Article IV Consultation* (IMF, 2019).
4. Ramachandran, V., Obado-Joel, J., Fatai, R., Masood, J. S. & Blessing, O. *The New Economy of Africa: Opportunities for Nigeria’s Emerging Technology Sector* (Center for Global Development, 2019).

PIUS UTO MI EKPEI/AFP VIA GETTY

Books & arts



Monuments to resilience or collapse? The 800-year-old statues of Easter Island.

Panicking about societal collapse? Plunder the bookshelves

As civilization seems to be lurching towards a cliff edge, historical case studies are giving way to big data in authors' search for understanding. **By Laura Spinney.**

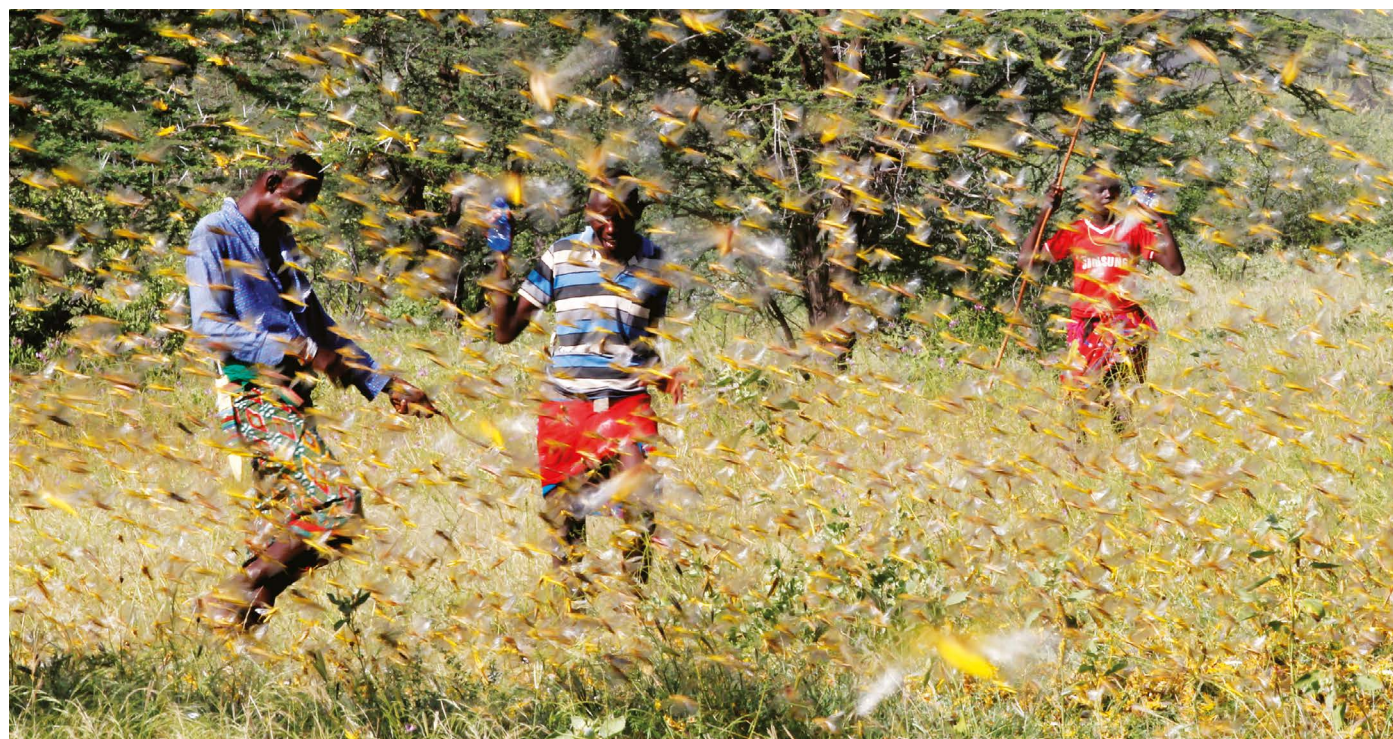
In case you missed it, the end is nigh. Ever since Jared Diamond published his hugely popular 2005 work *Collapse*, books on the same theme have been arriving with the frequency of palace coups in the late Roman Empire. Clearly, their authors are responding to a universal preoccupation with climate change, as well as to growing financial and political instability and a sense that civilization is lurching towards a cliff edge. Mention

is also made of how big-data tools are shedding new light on historical questions. But do these books have anything useful to share? Any

“What if collapse could usher in not only a renewed world, but a better one?”

actionable points besides that on my coffee mug: “Now panic and freak out”?

The newest is *Before the Collapse*. In it, energy specialist Ugo Bardi urges us not to resist collapse, which is how the Universe tries “to get rid of the old to make space for the new”. Similarly, Diamond’s 2019 book *Upheaval* suggested that a collapse is an opportunity for self-appraisal, after which a society can use its ingenuity to find solutions.



NJERINWANGI/REUTERS

East Africa is currently battling the biggest locust swarms in decades.

Both writers seem to accept that collapse is inevitable, but they take very different approaches to analysing it. Diamond zooms in to glean lessons from historical case studies; Bardi zooms out to view societies as complex dynamic systems that behave cyclically. Numerous books published in the past few decades chart how research has shifted from Diamond's approach to Bardi's.

Robust debate

Questioning Collapse, a 2009 collection of essays edited by archaeologists Patricia McAnany and Norman Yoffee, took Diamond to task for cherry-picking to spin a good yarn, for example in blaming such iconic societal failures as the population crash of Easter Island on its people's destruction of their own environment. The story is not so simple, the authors argue. The Indigenous Rapa Nui society weathered a string of environmental crises – very few of its own making – yet thrived until the first Europeans arrived. Likewise, is it reasonable to claim that Mayan society collapsed around the ninth century, given that seven million people living in and around Central America speak Mayan languages today? These cases might be better viewed, say McAnany and Yoffee, as lessons in resilience.

Scholars have long warned against peering down the 'retrospectroscope' at apparently neat examples of what not to do. In his influential 1988 *The Collapse of Complex Societies*, archaeologist Joseph Tainter argues that collapse – in the sense of the complete obliteration of a political system and its associated culture – is rare. Even the worst

cases are usually better described as rapid loss of complexity, with remnants of the old society living on in what rises from the ashes. After the 'fall' of Rome in the fifth century, for example, successor states took more than 1,000 years

Collapse: How Societies Choose to Fail or Succeed

Jared Diamond Viking (2005)

Before the Collapse: A Guide to the Other Side of Growth

Ugo Bardi Springer (2020)

Upheaval: Turning Points for Nations in Crisis

Jared Diamond Little Brown (2019)

Questioning Collapse: Human Resilience, Ecological Vulnerability, and the Aftermath of Empire

Edited by Patricia A. McAnany & Norman Yoffee Cambridge Univ. Press (2009)

The Collapse of Complex Societies

Joseph Tainter Cambridge Univ. Press (1988)

Understanding Collapse: Ancient History and Modern Myths

Guy D. Middleton Cambridge Univ. Press (2017)

Why the West Rules – for Now: The Patterns of History, and What They Reveal About the Future

Ian Morris Farrar, Straus and Giroux (2010)

War and Peace and War: The Rise and Fall of Empires

Peter Turchin Pi (2006)

Revolution and Rebellion in the Early Modern World

Jack Goldstone Univ. California Press (1991)

to achieve comparable economic and technological sophistication, but were always recognizably the empire's offspring.

Nevertheless, societies do go through rocky patches, from which some emerge transformed. It's not surprising that scholars should want to understand why. In his thoughtful *Understanding Collapse* (2017), archaeologist Guy Middleton surveys more than 40 theories of collapse – including Diamond's – and concludes that the cause is almost always identified as external to the society. Perennial favourites include climate change and barbarian invasions – or, in the Hollywood version, alien lizards. The theories say more about the theorists and their times, Middleton argues, than about the true causes of collapse.

Under strain

The pressing question, Tainter told a workshop on collapse at Princeton University in New Jersey last April, is why can a society withstand repeated external blows – until one day it cannot? For him, a society fails when it is no longer able to adapt to diminishing returns on innovation: when it can't afford the bureaucracy required to run it, say. In *Why the West Rules – For Now* (2010), historian Ian Morris proposes a twist on this, namely that the key to a society's success lies in its ability to capture energy – by extracting it from the ground, for example, or from nuclear fission once fossil fuels have run out. By contrast, Peter Turchin, author of the 2006 *War and Peace and War*, suggests that collapse is what happens when a society stops being able to deal with the strains caused by population

growth, leading to inequality and strife.

Turchin has been compared to Hari Seldon, science-fiction writer Isaac Asimov's "psycho-historian", who studies the past to statistically predict the future. He belongs to a new breed of scientific historian taking a big-data approach, and argues – controversially – that societal spasms are cyclic. This idea itself comes and goes: the ancient Greeks took the cyclic nature of history for granted, but it has been unfashionable since the Enlightenment. Today, we tend to have a linear concept of progress, in which life generally improves for most people over the long term. Works such as Turchin's see this trend as superimposed on an inherent cyclicity in the evolution of societies.

Reboot cycle

This raises the question of whether collapse is essential to renewal. Without winter, can you have spring? Bardi says no. Whether you think this good or bad depends partly on your point of view. The mass extinction 66 million years ago was bad for dinosaurs, but good for mammals, sociologist Miguel Centeno observed at the Princeton workshop, which he convened. But if collapse could usher in not only a renewed world, but a better one, shouldn't we dinosaurs embrace it?

For Turchin and Jack Goldstone – on whose work on the demographic forces shaping history Turchin builds – this is good advice only if you understand what causes collapse. Then it might be possible to make the transition less violent or disruptive. Goldstone rigorously dissected upheaval in the sixteenth to the nineteenth centuries in his 1991 book *Revolution and Rebellion in the Early Modern World*. This convinced him that revolution is an inappropriate response to societal tensions, usually leading to tyranny. Solutions have come instead from deep, meaningful reform. Yet the idea that revolution removes obstacles to progress has "deluded literally billions of people", he argues.

An interdisciplinary community of researchers is now searching for patterns that have defined collapse throughout history, to determine what might be an appropriate response. If we can't and shouldn't prevent a future crisis, could we at least soften it – perhaps with the help of new technologies – so that renewal happens, but less is lost and fewer people suffer? Even if the mind-boggling complexity of human societies makes this a pipe dream, as some argue, it seems a sounder approach than sparring over case studies that might not have constituted collapse at all. Speaking as a dinosaur, whose only alternative is to panic and freak out, I'll take it.

Laura Spinney is a science writer based in Paris. Her most recent book is *Pale Rider: The Spanish Flu of 1918 and How it Changed the World*. e-mail: lfspinney@gmail.com

Books in brief

CURED

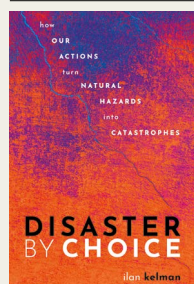


The Life-Changing Science of Spontaneous Healing
JEFFREY REDIGER, MD

Cured

Jeffrey Rediger Flatiron (2020)

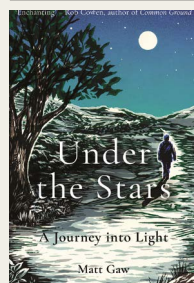
An experienced physician who is also a skilled, driven and compassionate writer is a winning combination. This pioneering book by psychiatrist Jeffrey Rediger analyses unexplained spontaneous recoveries from potentially fatal medical conditions, including cancer. From interviewing patients over nearly two decades, Rediger concludes that each recovery was "unique" and only partially explicable, but that all provide evidence of "a powerful link" between our identities and our immune systems.



Disaster by Choice

Ilan Kelman Oxford Univ. Press (2020)

Human choices cause disasters, but can also prevent them, argues Ilan Kelman in this grimly informative history. A specialist in disasters and health, he surveys earthquakes, epidemics, floods and more in a range of countries. Thus, in 2010, a magnitude-7.1 earthquake near Christchurch, New Zealand, caused not a single death. The same year, a magnitude-7.0 quake in Haiti caused at least 100,000 fatalities and a cholera outbreak – because of poor buildings and health care. Scientific foresight and political will are always key to resilience.



Under the Stars

Matt Gaw Elliott & Thompson (2020)

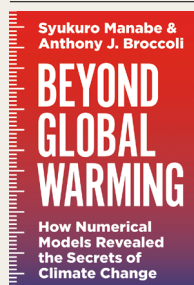
The Milky Way is invisible to 77% of today's UK population because of artificial light, notes naturalist and journalist Matt Gaw: "Many adults and children, my own included, have never seen it." Such thoughts inspired this poetically written but scientifically grounded study of darkness and its effect on humans and wildlife. Gaw describes night wanderings on English beaches, across Dartmoor and in central London. On the Scottish island Coll, a Dark Sky Community without a single street light, his children were entranced by the stars.



Stealth

Peter Westwick Oxford Univ. Press (2020)

In 1961, Dwight Eisenhower warned in his last address as US president that the "military-industrial complex" must be checked for the sake of "security and liberty". Historian Peter Westwick is more positive in his incisive narrative of the top-secret 1970s invention and construction of the stealth plane F-117. Nearly invisible to Soviet-designed radar, it was used to crucial effect in the 1991 Gulf War. Westwick argues that it offered an alternative to nuclear weapons, but admits that "to defend American liberties, aerospace engineers gave up civil liberties".



Beyond Global Warming

Syukuro Manabe & Anthony J. Broccoli Princeton Univ. Press (2020)

The first global climate model, developed in 1896 by chemist Svante Arrhenius, included the warming effect of atmospheric carbon dioxide. In the 1960s, meteorologist Syukuro Manabe pioneered computer simulation of climate change. Manabe's book written with atmospheric scientist Anthony Broccoli has evolved from his lecture notes, with chapters on, for example, general circulation models. Although technical, it should prove useful to those wishing to understand global warming's future impact. **Andrew Robinson**

Louis Nirenberg

(1925–2020)

Mathematician who transformed the study of partial differential equations.

After the Second World War, mathematics in the United States flourished owing to a convergence of interests. Mathematicians had shown their worth to military and industry patrons, who underwrote far-reaching empires of theories and people, including the consummate problem-solver Louis Nirenberg.

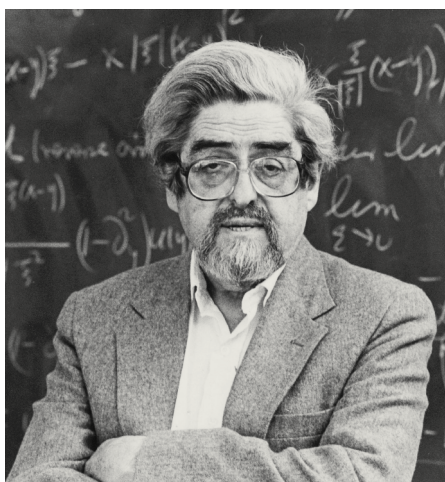
One of the world's most cited and productive mathematicians, Nirenberg was also among the most collaborative. His work continued to make waves until he was well into his eighties, and reshaped how mathematicians understand and study dynamical systems, from cells to markets. Winning the 2015 Abel Prize (shared with John Nash, made famous by the 2001 film *A Beautiful Mind*) was just a bookend to a fêted career. He died on 26 January, aged 94.

Nirenberg spent an illustrious seven decades at New York University (NYU), a realization of the discipline's post-war entanglements and their scholarly rewards. He joyfully nurtured people and ideas, skating above emerging distinctions between pure and applied mathematics.

Nirenberg transformed the field of partial differential equations (PDE), which explores what can be known about mathematical functions from studying how their variations along different dimensions relate to each other. Emerging from eighteenth-century mathematical physics, PDE became a centrepiece of a vast range of theoretical and applied subjects, from telecommunications and nuclear physics to debates about the nature of numbers. One famous and still-unresolved question in which Nirenberg's insights have been significant asks whether the equations governing the movement of water from a given initial state are always compatible with a smooth flow.

A virtuoso of approximation, Nirenberg was renowned for manipulating inequalities that govern the properties of unknown functions. Fellow mathematicians found his perspectives and methods strikingly lucid. His works catalysed large bodies of research, from general relativity to biology.

Raised in a Yiddish-speaking family in Montreal, Canada, Nirenberg acquired a taste for mathematical puzzles from his Hebrew tutor. After completing his undergraduate degree in mathematics and physics at McGill University in Montreal in 1945, Nirenberg joined his friend Sarah Courant at the nearby



National Research Council, contributing to research on atomic weapons. On the advice of Sarah's father-in-law, Richard Courant, a leading mathematician at NYU, Nirenberg did a master's in mathematics at the university. He remained there for the rest of his career, heading the Courant Institute of Mathematical Sciences from 1970 to 1972.

Nirenberg trained with a who's who of twentieth-century mathematics, including his PhD supervisor James Stoker and his mentor Kurt Friedrichs as well as visiting scholars from across Europe, the Soviet Union and the Americas. With fellow students Peter Lax and Cathleen Morawetz, he climbed the ranks to professor.

Courant had courted federal contracts and support during the Second World War to lay the groundwork for his institute. Expansive budgets from sources including the Office of Naval Research supported Nirenberg's research on elliptic equations (with applications from fluid dynamics to finance) and pseudo-differential operators (a foundation for an enormous variety of approaches in modern physics). One-quarter of his publications, including his first four in 1953, were in the institute's own journal, *Communications on Pure and Applied Mathematics*.

Nirenberg considered the world's mathematicians to be "one big family", and found inspiration in visiting and hosting colleagues from around the world. His first extended overseas research trip, in 1951–52, took him to Zurich, Switzerland, where he wrote up results from his thesis and attended lectures from stars of Courant's generation. In 1963, he took part in a landmark symposium on

PDE in Novosibirsk, which helped to redefine the relationship between the Soviet Union and the United States. There, he forged close friendships in an environment he compared to a voyage at sea. A later geopolitically significant trip took him to China toward the end of the Cultural Revolution. After being assigned a PhD thesis in Italian as the subject for a term paper during his graduate studies, he developed a lifelong affinity for Italy.

Nirenberg was known for using methods in their most fruitful generality. "I have made a living off the maximum principle," he quipped, referring to a fundamental technique for establishing inequalities in PDE. He demonstrated its versatile potential to researchers in many fields. As a young man, he had worried about his ability to formulate original problems. Yet Nirenberg gained a reputation for his exceptional insight and taste as a poser of problems that stretched the limits of research in mathematics and beyond.

His career awards included the first Crafoord Prize in 1982 and the first Chern Medal of the International Mathematical Union in 2010. Although he knew Nash, their Abel Prize recognized PDE work from separate parts of the firmament.

A famously congenial collaborator, Nirenberg co-authored papers to an extent unusual in mathematics. Some collaborations took place entirely by post, including the only work he published with his lifelong colleague Lax, conducted while Nirenberg was in Japan. Other collaborations – including with the 46 doctoral students he supervised – involved extended dialogues in front of a blackboard or while walking to a restaurant, as he digested new ideas in company.

Brit Shields works on the cultural history of twentieth-century mathematics; she is senior lecturer in the School of Engineering and Applied Science at the University of Pennsylvania in Philadelphia and a part-time programme administrator at the Courant Institute of Mathematical Sciences, where she knew Nirenberg. **Michael J. Barany** is lecturer in the history of science at the University of Edinburgh, UK, where he studies the global history and culture of modern mathematics. Both had interviewed Nirenberg for their research.
e-mails: bshields@seas.upenn.edu;
m.barany@ed.ac.uk

Comment



GIORDANO CIPRIANI/GETTY

Coral in a mangrove swamp in the Raja Ampat Islands, Indonesia.

Set a global target for ecosystems

James E. M. Watson, David A. Keith, Bernardo B. N. Strassburg, Oscar Venter, Brooke Williams & Emily Nicholson

The conservation community must be able to track countries' progress in protecting wetlands, reefs, forests and more.

Next week, representatives of more than 190 nations are gathering in Rome to discuss how to halt the biodiversity crisis during this decade and beyond. Since 2010, targets for conserving species have shaped policy and galvanized efforts to halt species loss worldwide, as part of the Convention on Biological Diversity (CBD; see www.cbd.int/sp/targets). Yet no such targets exist for ecosystems – despite the

wealth of evidence showing that their health and functions are essential to the processes that maintain all life¹.

Targets that are specific, measurable, attainable, relevant and timely (SMART) are central to project planning and have proved to be effective in policies that seek to address global problems. For example, during the 1980s, a group of 20 nations agreed to set various limits on the production and consumption

of chlorofluorocarbons. This helped to guide the phase-out of these substances under the Montreal Protocol, which came into effect in 1989 (ref. 2).

It is now possible to establish a SMART target for ecosystems, as well as metrics to track progress in meeting that goal. Nations are no longer limited by a lack of knowledge or methods when it comes to ecosystem mapping and assessment (see 'Under pressure'). What's more, they can use a proven and standardized approach for ecosystem risk assessment: the Red List of Ecosystems protocol, which was adopted by the International Union for Conservation of Nature (IUCN) in 2014.

We urge those attending next week's meeting to place an ecosystem-based goal and target alongside species-based ones in their discussions. Nations have a chance to ensure that all of the world's remaining intact ecosystems are retained by 2030, that overall ecosystem area and integrity increase by 2050, and that all that fall below a level of degradation defined by the Red List of Ecosystems protocol are restored.

The ratification of an international target will compel governments to act. This is the only way to halt the decline of ecosystems.

Species and ecosystems targets

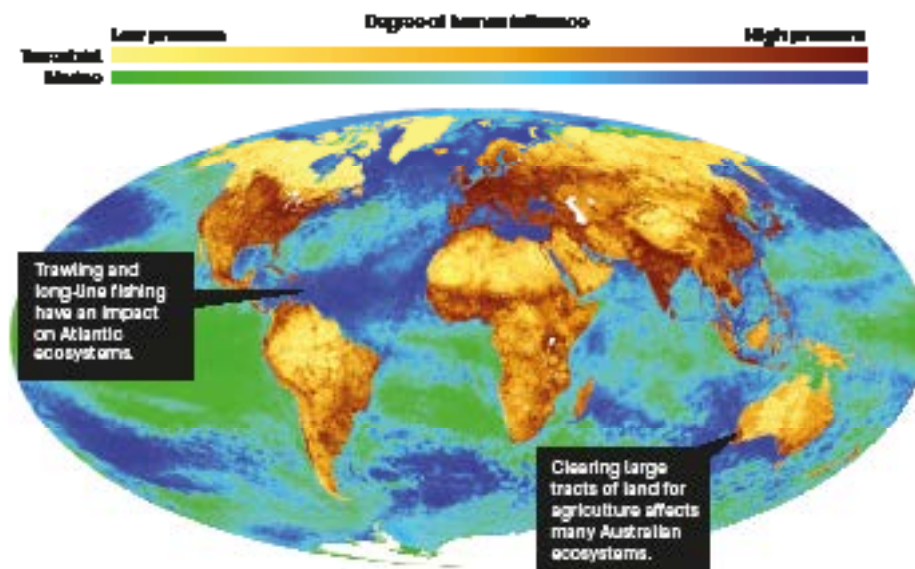
In 2010, the 193 nations that were parties to the CBD agreed to work together to prevent the extinction of known threatened species and improve their conservation status by 2020. They did this by ratifying Target 12 of the CBD 2011–20 strategic plan for biodiversity (see www.cbd.int/sp/targets).

Actions taken because of this and previous CBD targets have reduced the risk of extinction for many species, although direct links are hard to prove. For example, conservation efforts over the past 30 years have helped to cut the extinction rate of endangered birds by at least 40%, according to one analysis³. Previously endangered populations that are now growing include the Seychelles magpie-robin (*Copsychus sechellarum*) and a Brazilian parrot called Lear's macaw (*Anodorhynchus leari*).

Over the past decade, nations have been identifying and protecting the marine, terrestrial and freshwater sites that are of international importance to the conservation of vulnerable species. More than 16,000 of these 'key biodiversity areas' have now been identified worldwide (see go.nature.com/2xdtqb8). Government reports submitted to the CBD indicate that such areas are increasingly being protected⁴. One example is Itombwe

UNDER PRESSURE

Mapping of human pressures (from land use, development of infrastructure, mining and so on) reveals the degradation of natural systems across Earth.



Natural Reserve in the Democratic Republic of the Congo, which was formally established in 2016 to conserve several rare species, including the enigmatic Itombwe puddle frog (*Phrynobatrachus* sp.).

Such species-focused conservation activities are crucial. But they are not sufficient to sustain biodiversity and the benefits of nature to humanity.

Ecosystems, from the boreal forest and wetlands to coral reefs and mangroves, are more than the total of the plants and animals living in them⁵. Complex interactions between biological and physical systems drive processes that sustain all life. This includes production of clean water, regulation of air quality and climate through carbon sequestration and storage, soil formation, pollination and the

“Nations have a chance to ensure that all of the world’s remaining intact ecosystems are retained by 2030.”

production of food and wood for houses¹. Indeed, natural systems are key to dealing with the effects of climate change, as highlighted by a 2019 study⁶. It estimated that, between 2000 and 2013, the impact on carbon levels of losing intact tropical forests (including indirect effects such as reduced biodiversity and increased selective logging) might be six times greater than was originally proposed⁶.

Thanks to substantial advances in mapping and monitoring, scientists can now diagnose ecosystems’ defining features and the processes that threaten them^{5,7}. Take the demise of tidal flats revealed by satellite technology. Such mapping showed that coastal development and sea-level rise destroyed 16% of these ecosystems between 1984 and 2016. This has reduced storm protection and food provision for billions of people⁸. Remote sensing is similarly monitoring tropical forests⁹, ice cover¹⁰, coral reefs¹¹ and mangroves¹². For instance, at least 12% of the world’s mangroves were lost between 1996 and 2010 because of human activities¹³.

Pivotal to these efforts has been the development of the Red List of Ecosystems protocol, a set of criteria for identifying ecosystems that are most at risk of collapse¹⁴. It lays out how to define and map ecosystems, and enables systematic risk assessment using an array of indicators of extent and degradation.

So far, the Red List criteria have been used to assess more than 2,800 ecosystems in 100 countries across all continents¹⁵; 45% of those systems were found to be at risk of collapse (D.A.K., unpublished observation). These efforts could serve as a starting point for work towards an international target for conserving ecosystems.

Ecosystem-level conservation is already affecting decisions on resource use and management made by national governments, non-governmental organizations and industry¹⁵. For example, a 2017 assessment

Comment

of ecosystems in Colombia – Amazon rainforests, tropical dry forests, high Andean cloud forests, lowland savannah and other types – classified almost half (44%) as either ‘endangered’ or ‘critically endangered’, as defined by the Red List protocol¹⁶.

This results from human activities such as forest clearance for illegal coca crops, cattle ranching and mining. The finding has prompted the Colombian government to focus on the amount of land given protected area status, and to consider the restoration of critically endangered ecosystems.

In South Africa and Australia, businesses wanting to encroach on ecosystems that are classed as critically endangered or endangered must first conduct a full environmental impact assessment for their proposed project. Likewise, Finland’s first government-led systematic ecosystem assessment, completed in 2008, resulted in increased protection of threatened forest under the nation’s Environment Protection Act and Forest Act¹⁷.

In China, assessments of the rapid decline of tidal-flat ecosystems has catalysed efforts to better understand, manage and protect them. Tidal flats surrounding the Yellow Sea in east Asia support the migration of up to three million shorebirds and stabilize the coastline for more than 150 million people, also providing them with storm protection and food¹⁸. In July 2019, two of these important migratory sites were added to the United Nations Educational, Scientific and Cultural Organization (UNESCO) World Heritage List after being classified as endangered under the IUCN criteria.

Action and accountability

It is difficult to accurately assess progress towards conservation targets at the species level – a major constraint on their effectiveness. Monitoring of at-risk species is often infrequent and numbers fluctuate naturally from year to year. Such species also tend to be elusive. At the ecosystem level, a SMART target should therefore enable frequent tracking of ecosystems using remote sensing and modelling. This could result in more-transparent reporting of the status of Earth’s ecosystems, enhancing public awareness of their current trajectories and the consequences of their decline.

Any ecosystem target should set limits on degradation that mark the irreversible loss of key processes¹⁴. A target should also highlight the importance of conserving healthy ecosystems over restoring degraded ones. Such restoration is technologically and economically challenging and, as yet, there is no evidence that complete restoration of an ecosystem is even possible. Nevertheless, restoration has a key role in avoiding species extinctions and mitigating climate change, and should be part of an ecosystem goal.



RICARDO OLIVEIRA/AFP/GETTY

Human impacts such as overfishing have affected the Amazon River ecosystem in Brazil.

The Rome meeting is the second of three working-group meetings for negotiations leading up to a new set of biodiversity targets, which will replace those agreed in 2010. This 2030 global strategic plan for biodiversity will be formally established in October by the signatories to the CBD.

This year marks the implementation of the pledges made in the Paris climate agreement, and the United Nations Decade on Ecosystem Restoration begins in 2021. The launch of the 2030 strategic plan in October is an unprecedented opportunity – perhaps the last – for humanity to address multiple environmental problems at once. Whereas a species target forces nations to report on their progress only in relation to biodiversity, an ecosystem target would necessitate simultaneous reporting on wins across three fronts: biodiversity, climate change and sustainability (specifically, on the United Nations Sustainable Development Goals for human development and well-being).

World leaders must be held accountable for the current and future state of their countries’ ecosystems.

The authors

James E. M. Watson is a professor of conservation science at the University of Queensland, St Lucia, Australia, and director of the Science and Research Initiative at the Wildlife Conservation Society, Bronx, New York, USA. **David A. Keith** is a professor of botany at the University of New South Wales, Australia, and senior principal research scientist at the New South Wales Department of Planning, Infrastructure and Environment, Hurstville, Australia.

Bernardo B. N. Strassburg is a professor of sustainability science at the Pontifical Catholic University of Rio de Janeiro and executive director of the International Institute for Sustainability, Rio de Janeiro, Brazil. **Oscar Venter** is an associate professor at the Natural Resource and Environmental Studies Institute, University of Northern British Columbia, Prince George, Canada. **Brooke Williams** is a PhD candidate at the Centre for Biodiversity and Conservation Science, University of Queensland, St Lucia, Australia. **Emily Nicholson** is an associate professor at the Centre for Integrative Ecology, Deakin University, Burwood, Victoria, Australia. e-mail: jwatson@wcs.org

1. Díaz, S. et al. *Science* **366**, eaax3100 (2019).
2. Maxwell, S. L. et al. *Science* **347**, 1075–1076 (2015).
3. Monroe, M. J., Butchart, S. H. M., Mooers, A. O. & Bokma, F. *Biol. Lett.* **15**, 20190633 (2019).
4. UN Environment World Conservation Monitoring Centre, International Union for Conservation of Nature & Natural Geographic Society. *Protected Planet Report 2018: Tracking Progress Towards Global Targets for Protected Areas* (UNEP, 2018).
5. Keith, D. A. et al. *PLoS ONE* **8**, e62111 (2013).
6. Maxwell, S. L. et al. *Sci. Adv.* **5**, eaax2546 (2019).
7. Rowland, J. A. et al. *Conserv. Lett.* <https://doi.org/10.1111/conl.12680> (2019).
8. Murray, N. J. et al. *Nature* **565**, 222–225 (2019).
9. Hansen, M. C. et al. *Science* **342**, 850–853 (2013).
10. Parkinson, C. L. *Proc. Natl Acad. Sci. USA* **116**, 14414–14423 (2019).
11. Obura, D. O. et al. *Front. Mar. Sci.* **6**, 580 (2019).
12. Bunting, P. et al. *Remote Sens.* **10**, 1669 (2018).
13. Thomas, N. et al. *PLoS ONE* **12**, e0179302 (2017).
14. Keith, D. A. et al. *Conserv. Lett.* **8**, 214–226 (2015).
15. Bland, L. M. et al. *Conserv. Lett.* **12**, e12666 (2019).
16. Etter, A., Andrade, A., Saavedra, K., Amaya, P. & Arévalo, P. *Risk Assessment of Colombian Ecosystems: An Application of the Red List of Ecosystems Methodology (Vers. 2.0). Final Report*. (Pontificia Universidad Javeriana & Conservación Internacional-Colombia, 2017).
17. Kontula, T. & Raunio, A. *Biodivers. Conserv.* **18**, 3861–3876 (2009).
18. Murray, N. J., Ma, Z. & Fuller, R. A. *Austral Ecol.* **40**, 472–481 (2015).

Correspondence

Global solutions to prevent a pandemic

Investment in research must be fast-tracked if we are to tackle the new coronavirus disease, COVID-19. We need greater insight into the transmission, progression and epidemiology of this respiratory illness. We need to know the risk factors for infection, the role of asymptomatic or mild infection and the nature of 'super-spreaders'. We must determine disease seasonality and the viability of the virus in hot, humid environments, and improve estimates of death rates by age.

Research relevant to countries with weaker surveillance, lab facilities and health systems should be prioritized. In those regions, vaccine supply routes should not rely on refrigeration, and diagnostics should be available at the point of care. The World Health Organization is mapping such research and development priorities.

Social-science issues are important, too. These include how to communicate to the public what the options are for managing and preventing the disease, and how to tackle misconceptions and fear and avoid stigmatization. Community engagement and responsibility must be encouraged.

Charlotte H. Watts Government Department for International Development, London, UK.

Patrick Vallance Government Office for Science, London, UK.

Christopher J. M. Whitty Government Department of Health and Social Care, London, UK.
chris.whitty@dhsc.gov.uk

Careless virus names stoke sinophobia

The coronavirus that is currently causing severe respiratory illness worldwide has now been named SARS-CoV-2, and the disease is COVID-19. When the virus first emerged last December, it was generally described in medical journals as the '2019 novel coronavirus'. *Nature*, however, used 'China coronavirus' and 'Wuhan coronavirus'. Such interim terminology based on geographic characteristics is objectionable because it can stimulate prejudice and discrimination against Chinese people, fuelled internationally by fear spread through social media.

Although it is difficult and time-consuming to formally name diseases and viruses, it is essential that we methodically select no-harm names for them to make their way into human history. In 2015, the World Health Organization issued guidelines intended to minimize "unnecessary negative impact of disease names on trade, travel, tourism or animal welfare, and avoid causing offence to any cultural, social, national, regional, professional or ethnic groups". It asks scientists, journalists and health officials to use neutral, generic terms when referring to new human infectious diseases.

Lele Shu University of California, Davis, California, USA.
lele.shu@gmail.com

Editor's note: *Nature* has stopped referring to the 2019 novel coronavirus (SARS-CoV-2) as the Wuhan or China virus, for the reasons cited in the Correspondence. The names that appeared in earlier headlines were used to reflect the situation as it was understood at the time.

No special code for disaster research

As directors of the University of Delaware's Disaster Research Center, we disagree with the call by J. C. Gaillard and Lori Peek for a code of conduct for disaster-zone research (*Nature* 575, 440–442; 2019).

In our view, such a customized code would be likely to create a compliance morass out of all proportion to any ostensible harm. For example, the authors apply too broad a brush in referring to 'communities' and 'local priorities'. Communities are characterized by politics, power differences and stakeholders clamouring for attention. The authors suggest that research should align with community priorities. But rarely is there a single local priority, so whose priorities should take precedence, and why? Those priorities might even recreate the conditions that led to the disaster, or further marginalize other voices.

A disaster zone is not easy to define. The whole of Japan was affected by the 2011 Tohoku earthquake and tsunami, for example – even areas that were not physically hit. And, contrary to the authors' implication, there is no evidence that ethical concerns in post-disaster research are more severe than in other research involving human participants.

Such research can be done badly if, for example, the researcher has not properly reviewed the vast literature on quick-response best practice. Imposing criteria set by the United Nations would not prevent that. Dissemination and refinement of best practices remain the most crucial goals.

James Kendra, Tricia Wachtendorf University of Delaware, Newark, Delaware, USA.
jmkendra@udel.edu

Authorship: include citizen scientists

In our view, protocols for academic authorship need to adapt to acknowledge those members of the public who are increasingly engaging in important collaborations with researchers. These citizen scientists, who might include naturalists, farmers or Indigenous communities, rarely meet rigid journal-imposed criteria for authorship (see, for example, go.nature.com/2urkbrp). Consequently, protocols designed to stamp out ethical breaches, such as ghost authorship and conflicts of interest, exclude contributors who are not professional scientists.

Providing due credit is a core tenet of scientific ethics, and citizen scientists are pivotal to research projects and to the resulting publications.

Creating group co-authorships for cohorts of citizen scientists would credit them under a collective identity (see, for example, G. Ward-Fear *et al.* *Trends Ecol. Evol.* <http://doi.org/ggd6v7>; 2019). Furthermore, citizen scientists can play a crucial part in the uptake of scientific understanding by the general public.

Georgia Ward-Fear* Macquarie University, Sydney, Australia.
georgia.ward-fear@mq.edu.au

*On behalf of 4 correspondents (see go.nature.com/37kr9q5).

HOW TO SUBMIT

Correspondence may be submitted to correspondence@nature.com after consulting the author guidelines and section policies at go.nature.com/cmchno.

News & views

Epigenetics

How to silence an X chromosome

Jackson B. Trotman & J. Mauro Calabrese

The non-coding RNA *Xist* has been shown to enlist the SPEN protein to recruit a team of protein complexes – initiating the process that prevents transcription of one of the two X chromosomes found in female mammalian cells. **See p.455**

Female mammals have two X chromosomes, whereas males have only one. A remarkable solution has therefore evolved to prevent a gross imbalance in gene expression occurring between the sexes: in every cell that has two X chromosomes, one entire X chromosome is ‘silenced’ to prevent RNA from being transcribed from it. This process is called X-chromosome inactivation (XCI) and initiates early in the development of female embryos. Once complete, XCI is essentially stable for life¹ – thus, by extension, a human X chromosome can be propagated in the silenced state for more than 100 years.

XCI has become a paradigm for epigenetic processes – those in which DNA and associated proteins are modified to alter gene expression – and has been intensively studied for decades. For the past 25 years, much of this research has centred on a long non-coding RNA (lncRNA) called *Xist*, which is needed to orchestrate XCI. However, the details of *Xist*’s silencing mechanism have been elusive. Dossin *et al.*² report a stunning series of experiments on page 455 that reveal how *Xist* silences genes by partnering with a protein called SPEN.

Xist is expressed exclusively from the X chromosome that will be inactivated, where it spreads locally and silences nearly every gene on the chromosome by associating with an array of proteins. For example, *Xist* engages the Polycomb protein complexes (which modify the histone proteins that package DNA into a condensed form called chromatin) to maintain gene silencing on the inactivated X chromosome^{3,4}. Although this maintenance function is well documented, how *Xist* silences active genes in the first place has remained a mystery – in part because the majority of *Xist*’s protein partners were unknown. But in 2015, a series of studies^{5–9} revealed a comprehensive

set of proteins involved in XCI. These screens all identified SPEN as a *Xist*-binding protein that is essential for XCI.

SPEN belongs to an evolutionarily conserved family of RNA-binding proteins that have been implicated in transcriptional silencing and, curiously, RNA processing in both animals and plants¹⁰. To interrogate SPEN’s role in XCI, Dossin *et al.* first used a biological system known as an auxin-inducible degron to rapidly degrade SPEN in mouse embryonic stem cells. Consistent with a 2019 report¹¹, the authors observed that *Xist* is almost completely unable to silence genes along the X chromosome in the absence of SPEN. In an important first, the authors demonstrated that SPEN is required for successful XCI *in vivo* in mice. They also found that SPEN was needed

to dampen expression of ‘escapees’ – genes on the silenced X chromosome that partially evade XCI.

By observing fluorescently labelled molecules in living cells, Dossin *et al.* found that SPEN is recruited to the X chromosome as soon as *Xist* expression begins at the onset of XCI. SPEN contains four RNA-binding domains (called RRM) at its amino-terminal end and an evolutionarily conserved SPOC domain at its carboxy-terminal end. The authors found that, although RRM2–4 are required to bind *Xist*, the SPOC domain is the essential mediator of gene silencing. As suggested by previously reported experiments¹², forcing an interaction between *Xist* and the SPOC domain alone was enough to restore XCI in cells that lack SPEN.

It has been proposed^{7,13} that SPEN confers gene-silencing capabilities on *Xist* by recruiting and/or locally activating the enzyme HDAC3, which removes gene-activating acetyl groups from histones. However, HDAC3 accounts for only part of the gene silencing that occurs during the early stages of XCI¹³. To find other mechanisms by which SPEN might bring about silencing, Dossin *et al.* used a mass spectrometry technique to identify proteins that interact with the SPOC domain.

Confirming earlier work¹⁴, the authors found that SPEN’s SPOC domain interacts not only with HDAC3, but also with the associated co-repressor proteins NCOR1 and NCOR2 (also called SMRT), and with components of the nucleosome remodelling and deacetylase (NuRD) complex, all of which are epigenetic

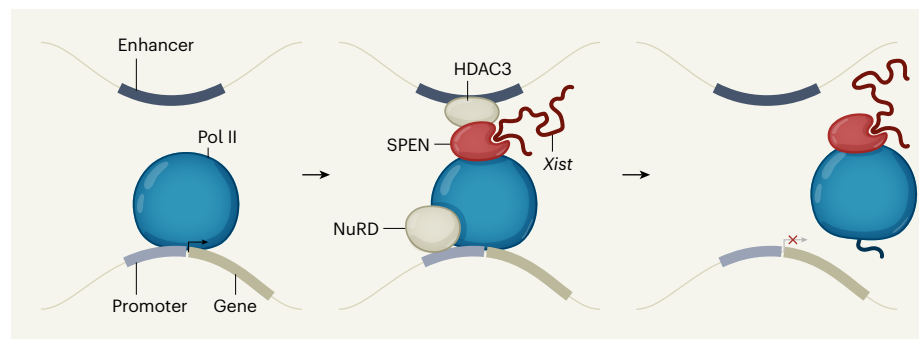


Figure 1 | Mechanism of gene silencing by SPEN. The long non-coding RNA *Xist* and its protein cofactor, SPEN, suppress (silence) gene expression in one of the two X chromosomes found in female mammalian cells. This is an essential process that prevents a gross imbalance in gene expression between males and females. Dossin and colleagues’ experiments² suggest that SPEN initiates this silencing mechanism by binding to active gene promoters (DNA sequences that initiate transcription) and enhancers (sequences that increase the likelihood of transcription). SPEN recognizes active promoters in part by interacting with constituents of the machinery used for gene transcription, including RNA polymerase II (Pol II, the enzyme that catalyses transcription). SPEN also recruits and/or locally activates the gene-inactivating protein HDAC3, and gene-silencing protein complexes such as the nucleosome remodelling and deacetylase (NuRD) complex. Once a gene has been silenced, SPEN disengages from its binding site, possibly displacing Pol II in the process.

silencers. Moreover, the authors observed that the SPOC domain interacts with parts of the machinery used for transcription and splicing (the process by which newly made RNA transcripts are turned into messenger RNA), including RNA polymerase II, the enzyme that catalyses transcription. Dossin and colleagues identified interactions with components of the *N*⁶-methyladenosine (m⁶A) methyltransferase complex, several of which have been linked to XCI^{6,11,15}. Accordingly, SPEN and its array of associated proteins might function like a molecular multi-tool to silence genes in various genomic contexts. Although much of SPEN's silencing function might derive from its interactions with known epigenetic silencers, its association with transcription and RNA-processing machineries leaves open the possibility that SPEN can also silence genes through another, as-yet-undefined mechanism.

Perhaps most strikingly, Dossin *et al.* adapted a technique called CUT&RUN to map the location of SPEN on an X chromosome that was being inactivated. This revealed that, shortly after *Xist* starts to be expressed, SPEN associates with active gene promoters and enhancers (DNA regions that initiate and increase the likelihood of transcription, respectively), but then disengages from these sites after it has silenced transcription. These discoveries imply that SPEN is part of a system that recruits silencing machinery specifically to transcriptionally active regulatory elements at the onset of XCI (Fig. 1). Whether this mechanism also requires chromatin modifications, RNA polymerase II, actively transcribed RNA or other factors should be addressed in the future. Another issue that should be investigated is why *Xist* isn't silenced by SPEN, given that a large amount of SPEN accumulates over the *Xist* gene.

SPEN binds to a region of *Xist* RNA called Repeat A, which is required to initiate gene silencing^{5,8,16}. Because deleting the *Spn* gene largely mirrors the effects of deleting Repeat A (ref. 11), SPEN seems to be responsible for most of Repeat A's silencing ability. However, Repeat A also binds to other proteins, including those that normally promote splicing, as well as to RBM15 and RBM15B, SPEN's SPOC-domain-containing cousins^{5,15,17}. Therefore, it is now crucial to determine how these proteins might compete or cooperate with SPEN to initiate gene silencing. Moreover, deletion of Repeat A drastically reduces levels of the *Xist* RNA itself⁸, and, in certain contexts, deletion of SPEN similarly reduces levels of *Xist*¹¹. How Repeat A is required for the production of *Xist*, and how its role in *Xist* production relates to its ability to initiate silencing, are key questions for the future.

For decades, *Xist* has served as a leading example of RNA's role in regulating gene expression. Most notably, *Xist* was one of the

first mammalian RNAs shown to be involved in Polycomb-mediated silencing^{3,4}. It therefore seems appropriate that, by studying this RNA, Dossin *et al.* might have uncovered a new and fundamental aspect of gene regulation – the transient recruitment of SPEN to regulatory elements by RNAs, or even by proteins, which could be a general mechanism for silencing transcription throughout the mammalian genome.

Jackson B. Trotman and **J. Mauro Calabrese** are in the Department of Pharmacology and the Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA.
e-mails: jbt@unc.edu; jmcablabr@med.unc.edu

1. Vallot, C., Ouimette, J.-F. & Rougeulle, C. *BioEssays* **38**, 869–880 (2016).
2. Dossin, F. *et al.* *Nature* **578**, 455–460 (2020).
3. Silva, J. *et al.* *Dev. Cell* **4**, 481–495 (2003).
4. Plath, K. *et al.* *Science* **300**, 131–135 (2003).
5. Chu, C. *et al.* *Cell* **161**, 404–416 (2015).
6. Moindrot, B. *et al.* *Cell Rep.* **12**, 562–572 (2015).
7. McHugh, C. A. *et al.* *Nature* **521**, 232–236 (2015).
8. Monfort, A. *et al.* *Cell Rep.* **12**, 554–561 (2015).
9. Minajigi, A. *et al.* *Science* **349**, aab2276 (2015).
10. Su, H., Liu, Y. & Zhao, X. *Cancer Transl. Med.* **1**, 21–25 (2015).
11. Nesterova, T. B. *et al.* *Nature Commun.* **10**, 3129 (2019).
12. Ha, N. *et al.* *iScience* **8**, 1–14 (2018).
13. Zyllicz, J. J. *et al.* *Cell* **176**, 182–197 (2019).
14. Shi, Y. *et al.* *Genes Dev.* **15**, 1140–1151 (2001).
15. Patil, D. P. *et al.* *Nature* **537**, 369–373 (2016).
16. Wutz, A., Rasmussen, T. P. & Jaenisch, R. *Nature Genet.* **30**, 167–174 (2002).
17. Pintacuda, G. *et al.* *Mol. Cell* **68**, 955–969 (2017).
18. Hoki, Y. *et al.* *Development* **136**, 139–146 (2009).

This article was published online on 5 February 2020.

Materials science

Why surface roughness is similar at different scales

Astrid S. de Wijn

Most surfaces are rough at many length scales. Simulations show that this characteristic originates at the atomic level in metal-based materials when smooth blocks of these materials are compressed.

Almost all solid surfaces are rough. This roughness occurs at length scales that encompass 13 orders of magnitude – from the kilometre-scale peaks of mountains, down to atomic-scale bumps. Roughness seems to emerge regardless of what is done to a surface. Yet there is little understanding of how this roughness comes about, and especially why it is often self-affine, meaning that a surface looks similar on different length scales. Writing in *Science Advances*, Hinkle *et al.*¹ show that self-affine roughness has its origin at the atomic level.

As anyone who has ever slipped on a wet floor will have noticed, the roughness of surfaces can have a crucial role in practical situations. Smooth surfaces are slippery when wet, but are also easier to lubricate inside moving machinery than are rough surfaces. By contrast, we sand surfaces before painting them to make them rougher, and thereby to increase the adhesion of the paint. The effects of roughness are less straightforward in other situations: for example, the roughness of the surfaces of skis and snowboards affects their friction on snow differently depending on the temperature and humidity². Engineers have therefore developed many techniques to control surface roughness, such as grinding, polishing and so on. Hinkle and colleagues' results help us to understand

better how roughness emerges, and thus might provide new ideas for how to control it.

The authors carried out computational simulations of three materials: a single, perfect gold crystal, an alloy and a metallic glass. These materials have very different amounts and types of disorder, which means that roughness might be expected to develop through different mechanisms or to have different characteristics for each of them. Because the deformation of a material is likely to contribute to the formation of roughness, the researchers simulated the compression of flat blocks of these materials beyond their elastic limit – that is, at forces that cause irreversible (plastic) deformation. Because the length scales of the effects the researchers were looking for span several orders of magnitude, the simulations had to be quite large, containing tens of millions of atoms. Such simulations are computationally extremely expensive.

Hinkle and colleagues investigated how fluctuations in the roughness produced in the simulations change when the size of the area being observed is increased. They observed that the roughness profiles of all three materials seem to obey a power law – that is, they do indeed display self-affine scaling, over nearly two orders of magnitude (from about 1 nanometre up to the size of their

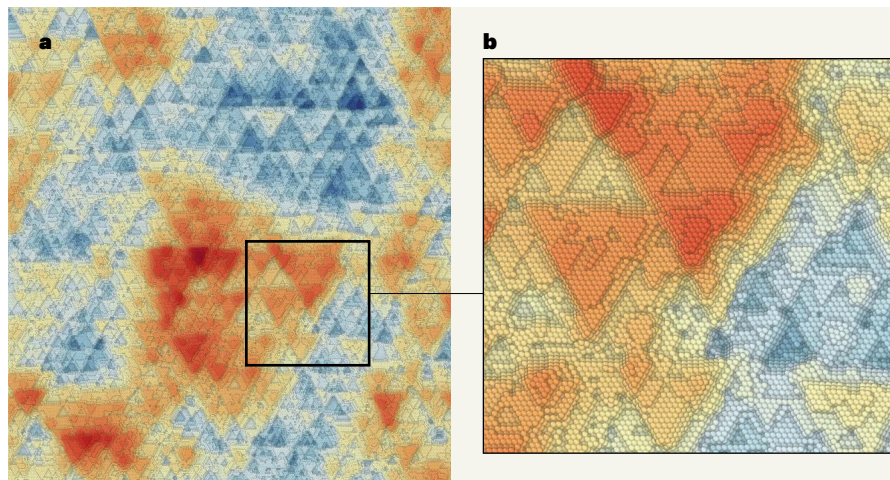


Figure 1 | Roughness on a simulated gold surface. Hinkle *et al.*¹ carried out molecular-dynamics simulations of tens of millions of atoms in smooth blocks of three materials, including gold (shown here), and observed how surface roughness develops when the blocks are compressed. Colours represent atomic positions perpendicular to the surface, measured relative to the surface's mean height: red indicates high topography; blue, low. The highest features are 8.8 nanometres above the lowest point on the surface. The authors found that roughness emerges that is similar across nearly two orders of magnitude of length scales. Similar triangular features and variation of topography are visible in **a** (a region 80 nm across) and **b** (a region of **a** expanded to four times its original size). The same is also true at magnifications of 8 and 64 (not shown).

simulation 'box', which was approximately 70–100 nm; Fig. 1).

In addition to simulating millions of atoms, the authors simulated a continuum model of compressive deformation in which the material is not treated as being composed of individual atoms, but as a continuous medium. In these simulations, there is no sign of self-affine roughness. The authors therefore conclude that the development of self-affine roughness is related to atomic-scale fluctuations in plastic flow that are missing from the continuum model.

Hinkle and colleagues' results are convincing across the observed length scales, but the scaling behaviour of the roughness will need to be demonstrated across three orders of magnitude to confirm that it truly obeys a power law. This will require the atomic simulations to be extended to even larger scales. Modelling techniques (see ref. 3, for example) are available at mesoscale lengths (which range from a few nanometres to several hundred micrometres), and provide a link between atomistic and continuum simulations. These approaches take flow into account in more detail than does the continuum model used by Hinkle *et al.*, and would allow for increased atomistic detail and fluctuations in simulations. This could help to provide the extra order of magnitude needed to convincingly show the power-law statistics of the roughness.

It remains to be seen how universal the reported behaviour is. All of the materials investigated by Hinkle and co-workers are based on metals. After undergoing plastic deformation, they are all homogeneous (there is only one type of solid phase in the material) but disordered, and the dynamics and energy

scales involved in atom displacement are all comparable. It would be interesting to see whether similar scaling behaviour emerges from the compression of other types of material that have different mechanisms of plasticity and deformation, such as polymers. If so, are the scaling exponents – the key scaling parameters in the power-law equation – the same for all materials? If roughness profiles can be extended to include one or more extra orders of magnitude, it would enable a reliable comparison of the scaling exponents.

Cancer

Loss of p53 protein strikes a nerve for tumour growth

Marco Napoli & Elsa R. Flores

Tumours often grow entangled among neurons, which makes the cancer difficult to treat. The finding that cancer cells hijack neighbouring neurons to promote tumour growth suggests new therapeutic targets. **See p.449**

Malignant tumours are a complex, yet organized, diverse ensemble of cells. Tumour cells are surrounded by other types of cell, which collectively form the tumour microenvironment. Components of this microenvironment include fibroblast cells, which can promote the growth and spread of tumours to distant sites, and immune cells. The latter have antitumour functions that are often suppressed

This, in turn, would help to determine whether these exponents vary with strain, deformation mechanisms, or even time.

Power-law behaviour is common in plastic deformation. For example, 'avalanches' of plastic deformation occur in metals⁴, and in fibrous materials a power law describes the size distribution of avalanches when these materials deform plastically under tensile stress⁵. Given that Hinkle *et al.* simulate the formation of rough surfaces in response to plastic deformation, and also observe scale-free roughness in the bulk of the modelled materials, it seems likely that there is a link between the development of self-affine roughness and the power-law behaviour of plastic deformation events – as the authors also note. It would therefore now be interesting to study the emergence of roughness in a more dynamic way, by investigating the formation of roughness features during compression, and relating the changes in the surface profile to plastic events.

Astrid S. de Wijn is in the Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Trondheim 7492, Norway.
e-mail: astrid.dewijn@ntnu.no

1. Hinkle, A. R., Nöhling, W. G., Leute, R., Junge, T. & Pastewka, L. *Sci. Adv.* **6**, eaax0847 (2020).
2. Colbeck, S. A. *A Review of the Processes that Control Snow Friction* (NTIS, 1992).
3. Bulatov, V., Abraham, F. F., Kubin, L., Devincere, B. & Yip, S. *Nature* **391**, 669–672 (1998).
4. Dahmen, K. A., Ben-Zion, Y. & Uhl, J. T. *Phys. Rev. Lett.* **102**, 175501 (2009).
5. Kloster, M., Hansen, A. & Hemmer, P. C. *Phys. Rev. E* **56**, 2615–2625 (1997).

This article was published online on 14 February 2020.

by cancer cells; indeed, therapies that boost such immune cells are revolutionizing the treatment of certain cancers. By contrast, the interactions between cancer cells and neurons in the tumour microenvironment are less well understood. On page 449, Amit *et al.*¹ reveal how tumours influence neurons to promote tumour growth, and show how this discovery might lead to new anticancer therapies.

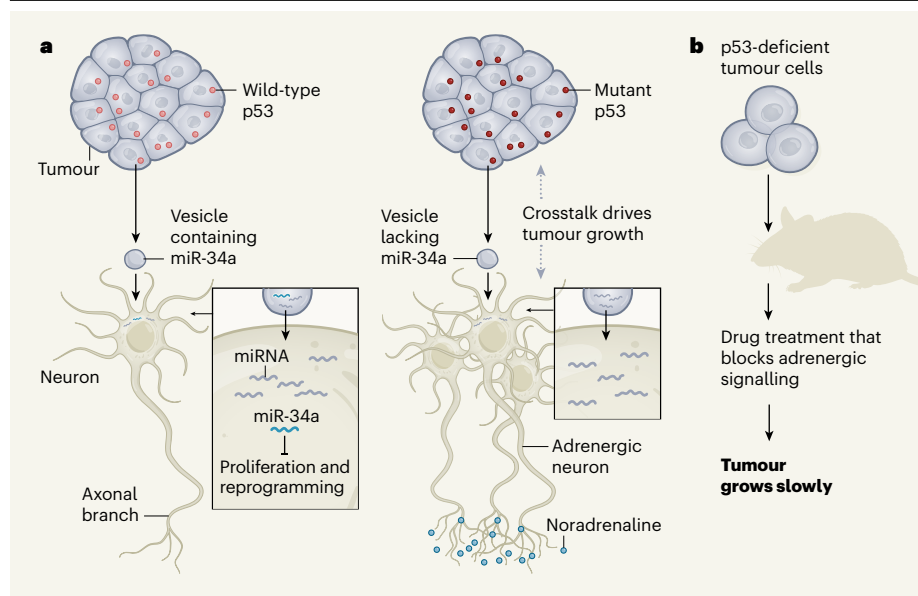


Figure 1 | Tumours manipulate neighbouring neurons to boost cancer growth. Amit *et al.*¹ analysed head and neck cancers using clinical data and mouse models. **a**, Tumour cells that expressed wild-type p53 protein released vesicles containing small RNA molecules called microRNAs (miRNAs) that were taken up by neighbouring neurons. An miRNA known as miR-34a blocks neuronal proliferation, and the neurons were maintained in their current state. By contrast, tumours that had a mutant version of p53 released vesicles that lacked miR-34a. In this case, neurons increased in number in the vicinity of the tumour, and these cells were reprogrammed as adrenergic neurons that express the molecule noradrenaline. These neurons had more axonal branches than did those near tumours that expressed wild-type p53. Interactions between adrenergic neurons and the tumour aided cancer growth. **b**, When mice received a transplant of p53-deficient tumour cells, treatment with a drug (carvedilol) that blocks adrenergic signalling pathways slowed tumour growth. This might provide a new therapeutic tool for targeting tumours that need neighbouring adrenergic neurons for their growth.

The interplay between cancer and neurons has negative clinical consequences for people with prostate tumours². Individuals who have a higher number of new neurons (in structures called nerve fibres) in the tumour microenvironment tend to have more-aggressive tumour features, such as further tumour growth and migration to distant sites, and a decrease in survival time².

Studies last year found that cancer cells and neurons can interact directly with each other through connections called synapses, that these connections aid the growth of brain tumours^{3–5} called gliomas, and that this interaction is associated with lethal cancer spread⁵. These and other findings^{3–6} contribute to a growing body of evidence that neurons are crucial components of the tumour microenvironment. However, what prompts the formation of neurons in the microenvironment had not been understood until now.

Amit and colleagues took on this challenge by focusing on tumours known as head and neck cancers, which can arise in the oral cavity. In humans, these tumours are often characterized by mutations that inactivate the gene *TP53*. This gene encodes a protein (p53) that functions as a tumour suppressor and that can modulate the tumour microenvironment⁷. By analysing four different mouse models of this

disease and data obtained from biopsies of people with head and neck cancer, the authors found that tumours with mutant versions of p53 have a higher number of associated newly formed neurons than do those with wild-type p53. Moreover, an increased number of such neurons correlated with a shorter survival time.

To try to determine whether cancer cells with mutant p53 might stimulate neurons to form, Amit *et al.* analysed the factors released

“The authors’ findings might have repercussions that reach beyond the field of cancer research.”

by human cancer cells that have mutant or wild-type p53. Both types of cell secreted vesicles that contained small RNA molecules called microRNAs (miRNAs). The vesicles in the two cell types were of a comparable number and size, but their contents differed (Fig. 1). Only the vesicles secreted from tumours with mutant p53 were devoid of an miRNA termed miR-34a, which is a tumour suppressor. When vesicles from tumours lacking p53 were injected into mice with tumours that had wild-type p53, the tumours with wild-type

p53 grew larger and had more surrounding neurons than normal, indicating that the contents of these vesicles drive the formation of new neurons. This is the first report showing that miR-34a, the main function of which is to keep in check the proliferation of normal and cancer cells⁸, is important in counteracting the formation of neurons in the tumour microenvironment.

Amit and colleagues analysed how these newly formed neurons promote tumour growth. The authors examined the neurons present in tumours with mutant and wild-type p53. Intriguingly, in the former set, the neurons (presumably including those already present in the area where the tumour formed) had undergone a functional change to become a type of neuron known as an adrenergic neuron – which uses the adrenergic signalling pathway and is activated in the ‘fight-or-flight’ response. This adrenergic feature (which has hallmarks including expression of the molecule noradrenaline) was crucial for sustaining cancer growth.

Interestingly, previous epidemiological analysis⁹ revealed that the use of the drug carvedilol, which blocks adrenergic signalling and is prescribed for conditions such as high blood pressure, is associated with a reduced risk of cancer onset. Now, Amit *et al.* raise the question of whether carvedilol’s anticancer properties might be due to its ability to target adrenergic neurons, given the effectiveness of the drug in treating mice with p53-deficient tumours (Fig. 1). The authors’ findings are of particular interest because these insights might offer a way to combat the tumour-driven formation of adrenergic neurons and to counteract their tumour-promoting effects. It will be important to establish whether adrenergic neurons’ contribution to tumour growth is limited to just head and neck cancers that have mutant p53, or whether this phenomenon could also be a feature of other types of tumour, as suggested by the epidemiological evidence for carvedilol use⁹.

Mutant versions of the gene encoding p53 are among the most common alterations in certain human cancers, occurring in approximately 60% of colon cancers, 50–80% of lung cancers and 95% of ovarian tumours¹⁰. Given the high prevalence of p53 abnormalities in cancer, numerous efforts have been made to design compounds that target mutant p53 to force it to act like wild-type p53, and promising results have been obtained in early-phase clinical trials of such drugs¹¹. It would be worth testing whether using both carvedilol and a drug that targets mutant p53 is more effective than either compound alone in treating these lethal forms of cancer.

Amit and colleagues’ discovery that the absence of functional p53 influences the formation of neighbouring neurons might have relevance for interpreting reports showing

that fluctuations in the levels of wild-type p53 are observed in nerve regeneration¹². Thus, the authors' findings might have repercussions that reach beyond the field of cancer research to regenerative medicine. Perhaps therapies that modulate the activity of p53 will have a future role in aiding the repair or regeneration of neurons, an outcome that would make a profound difference to the lives of people who have neurodegenerative diseases or other types of nerve injury.

Marco Napoli and **Elsa R. Flores** are in the Department of Molecular Oncology, H. Lee Moffitt Cancer Center and Research Institute,

Tampa, Florida 33612, USA.
e-mails: elsa.flores@moffitt.org;
marco.napoli@moffitt.org

1. Amit, M. et al. *Nature* **578**, 449–454 (2020).
2. Magnon, C. et al. *Science* **341**, 1236361 (2013).
3. Venkataramani, V. et al. *Nature* **573**, 532–538 (2019).
4. Venkatesh, H. S. et al. *Nature* **573**, 539–545 (2019).
5. Zeng, Q. et al. *Nature* **573**, 526–531 (2019).
6. Yu, K. et al. *Nature* **578**, 166–171 (2020).
7. Biegging, K. T. et al. *Nature Rev. Cancer* **14**, 359–370 (2014).
8. Zhang, L. et al. *J. Exp. Clin. Cancer Res.* **38**, 53 (2019).
9. Lin, C. S. et al. *Int. J. Cardiol.* **184**, 9–13 (2015).
10. Kandoth, C. et al. *Nature* **502**, 333–339 (2013).
11. Bykov, V. J. N. et al. *Nature Rev. Cancer* **18**, 89–102 (2018).
12. Krishnan, A. et al. *Eur. J. Neurosci.* **43**, 297–308 (2016).

This article was published online on 12 February 2020.

Atomic physics

Fundamental symmetry tested using antihydrogen

Randolf Pohl

The breaking of a property of nature called charge–parity–time symmetry might explain the observed lack of antimatter in the Universe. Scientists have now hunted for such symmetry breaking using the antimatter atom antihydrogen. **See p.375**

One of the greatest mysteries in modern physics is why the Universe seems to contain mostly matter and almost no antimatter. This observation could be explained if a property of nature called charge–parity–time (CPT) symmetry is violated. Under CPT symmetry, the physics of particles and their antiparticles is identical. A tiny violation of CPT symmetry during the Big Bang could, in principle, be responsible for the lack of antiparticles in the Universe. On page 375, the ALPHA Collaboration¹ reports high-precision spectroscopic measurements of antihydrogen – an atom comprising an antiproton and a positron (the antiparticle of an electron). The authors find that the gaps between energy levels in antihydrogen are in excellent agreement with those measured previously in ordinary hydrogen^{2–4}, placing strong constraints on potential CPT violation.

Tests of CPT symmetry using individual particles – such as neutral kaons⁵, positrons⁶ and antiprotons^{7,8} – have shown no sign of CPT violation. However, studies of antihydrogen might probe the influence of factors that were not explored in previous tests.

Hydrogen is the simplest atom, and its properties can be calculated with impressive precision. For more than a century, the study of this atom has been the driving force behind groundbreaking ideas about the structure of matter. The optical spectrum of hydrogen was

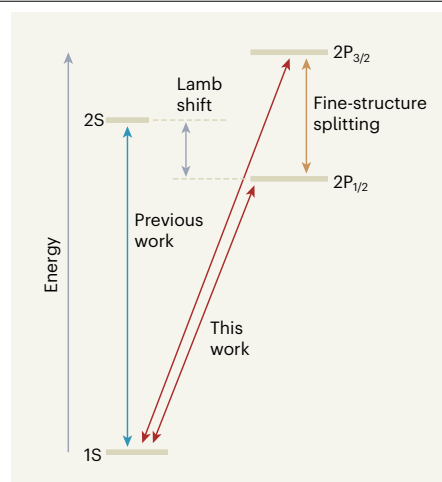


Figure 1 | Lowest-energy states of antihydrogen.

The ALPHA Collaboration¹ carried out high-precision spectroscopic measurements of antihydrogen – the antimatter counterpart of hydrogen. Specifically, the team determined the energy differences between the 1S ground state and the $2P_{1/2}$ and $2P_{3/2}$ excited states of antihydrogen. They used these results to estimate the fine-structure splitting (the $2P_{1/2}$ – $2P_{3/2}$ energy gap). They also combined their previous determination¹¹ of the energy gap between the 1S and 2S states with their current measurement of the 1S– $2P_{1/2}$ energy difference to infer the Lamb shift (the 2S– $2P_{1/2}$ energy gap). The authors found that all of these results are in agreement with the corresponding ones for ordinary hydrogen. (Drawing not to scale.)

measured with great accuracy in the 1880s, before being quantitatively explained in the 1910s. The structure of the atom was then at the heart of the formulation of quantum mechanics and in the generalization of this theory to relativistic (fast-moving) particles in the 1920s. And it was the unexpected discovery⁹ of an energy gap between the 2S and $2P_{1/2}$ excited states of hydrogen by the physicist Willis Lamb in 1947 that motivated the development of quantum electrodynamics – the theory that describes the interactions between particles and light.

This energy gap, known as the Lamb shift, exists in both hydrogen and antihydrogen. It originates mostly from quantum fluctuations, whereby particle–antiparticle pairs spontaneously emerge in empty space and then instantly annihilate each other. However, its magnitude is subtly affected by, for example, the charge radius (the spatial extent of the charge distribution) of the proton or antiproton, the weak nuclear force and, potentially, currently unknown phenomena that could be the source of the matter–antimatter asymmetry in the Universe.

The current work was carried out using the ALPHA experiment at CERN, Europe's particle-physics laboratory near Geneva, Switzerland. A facility called the Antiproton Decelerator delivers antiprotons to this experiment, with a source of radioactive sodium providing positrons. Every few minutes, 90,000 cold trapped antiprotons and 3 million positrons are mixed in a sophisticated charged-particle trap. This process yields about 20 cold antihydrogen atoms that are then confined to a neutral-atom trap made from superconducting magnets. These antihydrogen atoms can be stored¹⁰ for at least 60 hours, and production cycles can be repeated to obtain hundreds of such atoms.

The aim of the present study was to measure the energy differences between the 1S ground state and the $2P_{1/2}$ and $2P_{3/2}$ excited states of antihydrogen (Fig. 1). The ALPHA Collaboration used an approach called laser spectroscopy, which involved injecting pulses of laser light into the antihydrogen trap. This injection caused atoms to transition from the 1S state to the $2P_{1/2}$ or $2P_{3/2}$ state and to subsequently decay back to the 1S state. Atoms that ended up in a different magnetic substate of the 1S state from the one in which they started were expelled from the magnetic neutral-atom trap. These antihydrogen atoms then annihilated on contact with regular atoms in the walls of the ALPHA apparatus to produce particles called charged pions.

The ALPHA Collaboration plotted the number of observed charged pions as a function of the frequency of the laser light. They then used the positions of the two peaks in these plots to infer the 1S– $2P_{1/2}$ and 1S– $2P_{3/2}$ energy differences in antihydrogen.

These differences agree with the ones measured in ordinary hydrogen at the level of 16 parts per billion. The authors used their results to estimate the fine-structure splitting (the $2P_{1/2}$ – $2P_{3/2}$ energy difference) in antihydrogen, with an uncertainty of 0.5%. This value is again in good agreement with the one for ordinary hydrogen.

In 2018, the ALPHA Collaboration measured the energy gap between the 1S and 2S states in antihydrogen¹¹ to one part in 10^{12} . In the current work, the authors combined this result with their measurement of the 1S– $2P_{1/2}$ energy difference to provide an estimate of the Lamb shift in antihydrogen. This value has an uncertainty of 11% (or 3.3%, when the fine-structure splitting in ordinary hydrogen is used in the analysis).

Over the past few years, high-precision laser spectroscopy of antihydrogen has become possible, and the ALPHA Collaboration has achieved spectacular progress. An examination of several transitions in antihydrogen would enable targeted tests of CPT symmetry, quantum electrodynamics and the standard model of particle physics. For example, a measurement of the Lamb shift with an uncertainty of less than one part in 10^4 would allow the antiproton charge radius to be determined¹². Moreover, improved measurements of the energy gap between magnetic substates in antihydrogen would provide detailed information about the magnetic structure of the antiproton¹³.

The laser used for spectroscopy in the current work will, in the future, be used for cooling of antihydrogen by inducing $1S$ – $2P_{1/2}$ and $1S$ – $2P_{3/2}$ transitions. Such cooling would greatly improve the achievable precision in all spectroscopy experiments on antihydrogen. In addition, ultracold antihydrogen can be used to study the effect of gravity on these atoms¹⁴. Cold antihydrogen thus promises many cool results.

Randolf Pohl is at the Institute of Physics, University of Mainz, 55128 Mainz, Germany. e-mail: pohl@uni-mainz.de

1. The ALPHA Collaboration. *Nature* **578**, 375–380 (2020).
2. Hagley, E. W. & Pipkin, F. M. *Phys. Rev. Lett.* **72**, 1172–1175 (1994).
3. Parthey, C. G. et al. *Phys. Rev. Lett.* **107**, 203001 (2011).
4. Bezginov, N. et al. *Science* **365**, 1007–1012 (2019).
5. Schwingenheuer, B. et al. *Phys. Rev. Lett.* **74**, 4376–4379 (1995).
6. Dehmelt, H., Mittleman, R., Van Dyck, R. S. Jr & Schwinberg, P. *Phys. Rev. Lett.* **83**, 4694–4696 (1999).
7. Hori, M. et al. *Nature* **475**, 484–488 (2011).
8. Smorra, C. et al. *Nature* **550**, 371–374 (2017).
9. Lamb, W. E. Jr & Retherford, R. C. *Phys. Rev.* **72**, 241–243 (1947).
10. Capra, A. & ALPHA Collaboration. *Hyperfine Interact.* **240**, 9 (2019).
11. Ahmadi, M. et al. *Nature* **557**, 71–75 (2018).
12. Crivelli, P., Cooke, D. & Heiss, M. W. *Phys. Rev. D* **94**, 052008 (2016).
13. Zemach, A. C. *Phys. Rev.* **104**, 1771–1781 (1956).
14. The ALPHA Collaboration & Charman, A. E. *Nature Commun.* **4**, 1785 (2013).

In retrospect

30 years of the iron hypothesis of ice ages

Heather Stoll

In 1990, an oceanographer who had never worked on climate science proposed that ice-age cooling has been amplified by increased concentrations of iron in the sea – and instigated an explosion of research.

Thirty years ago this month, John Martin proposed a solution to one of the biggest mysteries of Earth's climate system: how was nearly one-third of the carbon dioxide in the atmosphere (about 200 gigatonnes of carbon) drawn into the ocean as the planet entered the most recent ice age, then stored for tens of thousands of years, and released again as the ice sheets melted? These large natural cycles in atmospheric CO_2 levels (Fig. 1a) were revealed in 1987 by an analysis of ancient air bubbles trapped in the first long ice cores taken from the Antarctic ice sheet¹. Martin recognized that iron was a key ingredient that could have transformed the surface ocean during glacial times. His landmark iron hypothesis², published in *Paleoceanography*, described a feedback mechanism linking climatic changes to iron supply, ocean fertility and carbon storage in the deep ocean.

Two hundred gigatonnes is a lot of carbon to periodically withdraw from and release to the atmosphere. In the 1980s, a handful of models (see ref. 3, for example) had shown that an increase in biomass production in polar ocean regions was the most effective process for removing so much atmospheric carbon. Photosynthetic organisms in the surface ocean convert CO_2 from the atmosphere into biomass, much of which is subsequently broken down into CO_2 again by other organisms and returned to the atmosphere. But part of the biomass sinks into the deep ocean, which therefore effectively serves as a large storage reservoir of dissolved CO_2 . This mechanism of CO_2 removal is called the biological pump.

However, biomass production requires not only CO_2 , but also other nutrients to build lipids, proteins and enzymes. Researchers were struggling to ascertain how the ocean's abundance of key nutrients, such as nitrates or phosphates, might have increased during glacial times to fuel a stronger biological pump.

Martin argued that iron is another nutrient that limits the biological pump. He suggested that the modern marine ecosystem

of the Southern Ocean around Antarctica is starved of iron, and therefore relatively low in biomass, despite having abundant nitrates and phosphates. But during glacial times, strong winds over cold, sparsely vegetated continents could have transported large amounts of iron-bearing dust into this ocean (Fig. 1b). Martin reasoned that this dust could have fertilized marine ecosystems and strengthened the biological pump, so that more carbon was transferred into the deep ocean, lowering atmospheric CO_2 levels.

Around the time of publication, evidence for high dust delivery during glacial periods had just emerged from studies of deep Antarctic ice cores⁴. But there were no reliable measurements of dissolved iron in the Southern Ocean that could confirm that its surface waters are iron-starved in modern times, or data supporting the proposal that delivery of iron-rich dust would make a difference to ocean productivity. It was clear, however, that large patches of the world's ocean had much lower quantities of biomass than would be expected on the basis of the concentrations of key nutrients such as nitrates and phosphates. But many researchers argued that this was due to natural overgrazing of algae by herbivores⁵.

The idea that modern algal growth is limited by iron availability had, in fact, been proposed⁶ in the 1930s, but had been incorrectly discounted by oceanographers – who had measured plenty of iron in seawater samples collected from the waters around their iron ships⁷. Martin was one of the first oceanographers to implement painstaking procedures to avoid the contamination of samples and to determine that iron concentrations in the north Pacific Ocean were extremely low⁷, certainly low enough to curtail biomass production.

Despite the initial scepticism that greeted the iron hypothesis, 12 separate experiments⁸ were carried out between 1993 and 2005 in which around 300–3,000 kilograms of dissolved iron were injected into small patches of the Southern Ocean, the equatorial Pacific

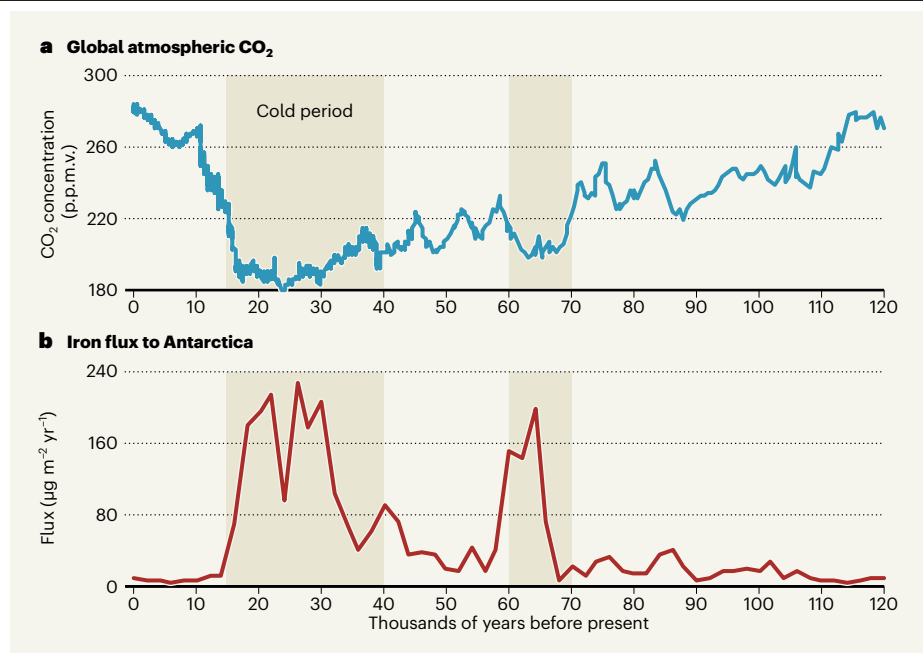


Figure 1 | The anti-correlated data that inspired the iron hypothesis. **a**, Measurements of air bubbles trapped in cores drilled from the Antarctic ice sheet show that atmospheric levels of carbon dioxide were significantly lower during the coldest periods (shaded regions) than during modern times (data from ref. 16; CO₂ concentrations are shown in parts per million by volume; p.p.m.v.). **b**, The ice-core records also reveal that more iron was transported to the Southern Ocean in wind-blown dust during the coldest periods than during warmer times (data from ref. 17; iron flux is measured in micrograms per square metre per year). In 1990, Martin² hypothesized that the increased levels of iron in the Southern Ocean during the coldest periods fertilized the growth of photosynthetic microorganisms in the surface Southern Ocean, which therefore produced more biomass from CO₂. This, in turn, would have increased the strength of the biological pump, a mechanism that sequesters some of the biomass (and the carbon within it) in the deep ocean. Martin proposed that the stronger biological pump explains why so much atmospheric CO₂ is drawn into the ocean during cold times.

Ocean and the north Pacific. The biomass of algae increased wherever iron was added, as biological production surged.

Unfortunately, Martin died mere months before the first of these experiments, and did not witness the ocean-scale confirmation of his hypothesis, nor the internationally coordinated campaign to measure iron geochemistry throughout the world's oceans⁹ – which confirmed iron limitation and revealed the intricate strategies used by marine ecosystems to acquire and recycle iron¹⁰.

Earth scientists also tried to test the iron hypothesis computationally using simple ocean models. They used the changes in the dust-accumulation rate recorded in ice cores as input to simulate changes in iron delivery to the Southern Ocean, and data from the experimental iron fertilizations to calculate how this iron could affect algal growth and the biological pump. Such models could reproduce the timing and magnitude of about half of the observed decrease in atmospheric CO₂ levels during glacial periods¹¹. Iron fertilization is therefore clearly an important process that causes atmospheric changes, but might not be the only one.

Finding data to prove that biological production had been higher during glacial

was a harder task – after all, the ecosystem during the most recent glacial period (about 20,000 years ago) is long dead. One possible solution was to extract cores from sediments piled on the sea floor, to see whether the mineral skeletons of algae accumulated faster during glacial times than in the modern era. However, the results were often ambiguous¹², for several reasons: many algae don't produce a preservable skeleton; numerous factors determine what proportion of biological remains is preserved on the sea floor; and the location of biological production changes through time as ocean fronts and sea-ice positions migrate.

Fortunately, Martin² and others¹³ had anticipated an alternative, global-scale test of the biological pump during glacial times. If more biomass reached the deep ocean during glacial, then deep-sea microorganisms would use up more oxygen as they consumed it, decreasing the concentration of oxygen in deep waters. Evidence of deep-ocean oxygen depletion would therefore be indicative of a strong biological pump.

Martin recognized that the presence of certain microfossils in glacial-age sediments meant that the deep ocean had not become completely devoid of oxygen during glacial.

But although this evidence crudely constrained estimates of the degree to which iron fertilization might have enhanced productivity during glacial, it could not be used to determine whether levels of deep-ocean oxygen were lower than during modern times. Since then, analysis of more-sensitive geochemical records indicates that the oxygen concentration in bottom waters did decrease during glacial times¹⁴. This provides the strongest confirmation yet of the large-scale accumulation of carbon in the deep ocean during glacial periods owing to a stronger biological pump.

Slower rates of mixing between the deep and shallow oceans could also have enhanced the biological pump during glacial. The latest generation of climate models in which the ocean and atmosphere are coupled can test the contribution of the multiple processes that could have resulted in a reduction in bottom-water oxygen levels. Such models indicate that mixing rates can account for only half of the observed deep-ocean storage of CO₂ during the glacial period, and that iron fertilization of the Southern Ocean is the major cause of the extra CO₂ storage observed¹⁵.

Martin concluded his paper by saying that iron availability “appears to have been a player” in strengthening the biological pump during glacial cycles, but that the size of its role remained to be determined. Thirty years later, the evidence convincingly shows that iron fertilization of the Southern Ocean was indeed a leading actor in this global-climate feedback.

Heather Stoll is in the Department of Earth Sciences, ETH Zurich, 8092 Zurich, Switzerland.
e-mail: heather.stoll@erdw.ethz.ch

1. Barnola, J. M., Raynaud, D., Korotkevich, Y. S. & Lorius, C. *Nature* **329**, 408–414 (1987).
2. Martin, J. H. *Paleoceanography* **5**, 1–13 (1990).
3. Sarmiento, J. L., Toggweiler, J. R. & Najjar, R. *Phil. Trans. R. Soc. Lond. A* **325**, 3–21 (1988).
4. De Angelis, M., Barkov, N. I. & Petrov, V. N. *Nature* **325**, 318–321 (1987).
5. Cullen, J. J. *Limnol. Oceanogr.* **36**, 1578–1599 (1991).
6. Hart, T. J. *Discov. Rep.* **VIII**, 1–268 (1934).
7. Gordon, R. M., Martin, J. H. & Knauer, G. A. *Nature* **299**, 611–612 (1982).
8. Boyd, P. W. et al. *Science* **315**, 612–617 (2007).
9. Anderson, R. F. & Henderson, G. M. *Oceanography* **18**, 76–79 (2005).
10. Tagliabue, A. et al. *Nature* **543**, 51–59 (2017).
11. Watson, A. J., Bakker, D. C. E., Ridgwell, A. J., Boyd, P. W. & Law, C. S. *Nature* **407**, 730–733 (2000).
12. Kohfeld, K. E., Le Quéré, C., Harrison, S. P. & Anderson, R. F. *Science* **308**, 74–78 (2005).
13. Boyle, E. A. *Nature* **331**, 55–56 (1988).
14. Jaccard, S. L., Galbraith, E. D., Martínez-García, A. & Anderson, R. F. *Nature* **530**, 207–210 (2016).
15. Yamamoto, A., Abe-Ouchi, A., Ohgaito, R., Ito, A. & Oka, A. *Clim. Past* **15**, 981–996 (2019).
16. Bereiter, B. et al. *Geophys. Res. Lett.* **42**, 542–549 (2015).
17. Wolff, E. W. et al. *Nature* **440**, 491–496 (2006).

This article was published online on 17 February 2020.

Transfer of ubiquitin protein is caught in the act

Raymond J. Deshaies & Nathan W. Pierce

A process termed ubiquitination mediates the regulated destruction of cellular proteins, thereby preventing disease or infection. Structural data now reveal how a crucial regulator of ubiquitination enzymes coordinates this process. **See p.461**

Many cellular functions that occur in eukaryotes (organisms whose cells contain a nucleus) are regulated by targeted protein destruction. This targeting is often achieved by a process called ubiquitination (or ubiquitylation), in which a protein selected for destruction is tagged with the protein ubiquitin. Ubiquitination is aided by enzymes known as E3 ligases, a subset of which are called cullin–RING ubiquitin ligases (CRLs)¹. CRLs help to transfer ubiquitin from an E2 conjugating enzyme, to which it is bound, onto the target protein¹. By default, CRLs are inactive, and they are activated when a protein called NEDD8 (which is similar in sequence to ubiquitin) becomes attached to the cullin subunit of the CRL^{2–5}. But how this activation happens has been a mystery. On page 461, Baek *et al.*⁶ report structural data obtained using a technique called cryo-electron microscopy (cryo-EM) that fills in some of the blanks.

CRLs contain a banana-shaped cullin subunit (one of five cullin proteins, CUL1 to CUL5). This binds (Fig. 1) at one end to a substrate-receptor subunit – which recruits the protein targeted for ubiquitination – and at the other end to what is termed a RING-finger protein, which is either RBX1 or RBX2 (refs 1,7). The RING-finger protein recruits a ubiquitin-attached E2 enzyme and stimulates the transfer of its ubiquitin to the target protein¹. Previous structural analysis demonstrated that the attachment of NEDD8 to CUL5 enhances the potential of RBX2 and its ubiquitin-bound E2 enzyme to move towards the region adjacent to the substrate receptor and its bound target protein³. However, that work used a truncated version of CUL5 bound to RBX2, and lacked both a target protein bound to the substrate receptor and a ubiquitin-attached E2 enzyme, thus leaving to the imagination the mechanism by which NEDD8 stimulates

ubiquitin transfer to the target protein.

Baek and colleagues therefore sought to capture a human NEDD8-attached CRL in the act of transferring ubiquitin to its target protein. To achieve this goal, the authors made a ‘tribrid’ molecule comprising three components. One component was a stretch of amino-acid residues derived from the protein IκB, which is a ubiquitination target that binds to a substrate receptor called β-TRCP. The second was an E2 enzyme termed UBE2D, and the third was ubiquitin. This tribrid provided a stable mimic of how the molecular components are arranged during the transition state, when ubiquitin is being transferred from the E2 enzyme to the target protein. Using cryo-EM, the authors obtained structural data for the complex that formed when the tribrid and β-TRCP assembled with the proteins CUL1, SKP1 and RBX1 (this complex is called CRL1^{β-TRCP}).

The type of structural information that can be obtained using X-ray crystallography is constrained by technical issues (packing forces in the crystals) that affect data collection. The cryo-EM approach taken by the authors avoids these constraints and enables multiple conformations of a structure to be obtained. The authors confirmed an earlier finding³ that CRL1^{β-TRCP} shows modest conformational flexibility in the absence of NEDD8, but that this flexibility increases when NEDD8 is attached. Furthermore, on addition of the tribrid, the ensemble of conformations converged to form one structure of a transition-state intermediate.

Baek and colleagues’ structural data are nothing short of spectacular. Previous work³ suggested that the active site of the E2 enzyme, where ubiquitin is transferred to the target protein, might not be fixed in location relative to the NEDD8-attached CRL because of the mobility of the RING-finger protein’s RING domain. The transition state presented by Baek *et al.* shows the precise 3D relationship of three modules that form the whole complex: a catalytic module, an activation module and a substrate-scaffolding module. The catalytic module comprises ubiquitin-bound UBE2D and the RING domain of RBX1, and this module moves when NEDD8 becomes attached to CUL1. The activation module consists of a mobile domain in CUL1 called the WHB domain, to which NEDD8 attaches. The substrate-scaffolding module includes β-TRCP and portions of CUL1 and RBX1 that have a fixed spatial relationship to β-TRCP and IκB.

In Baek and colleagues’ proposed activated structure, the catalytic module projects directly towards the substrate-scaffolding module, such that UBE2D touches β-TRCP (Fig. 1). The activation module coordinates the architecture of the transition state, with NEDD8 forming multiple contacts between UBE2D in the catalytic module and CUL1 in the substrate-scaffolding module. These

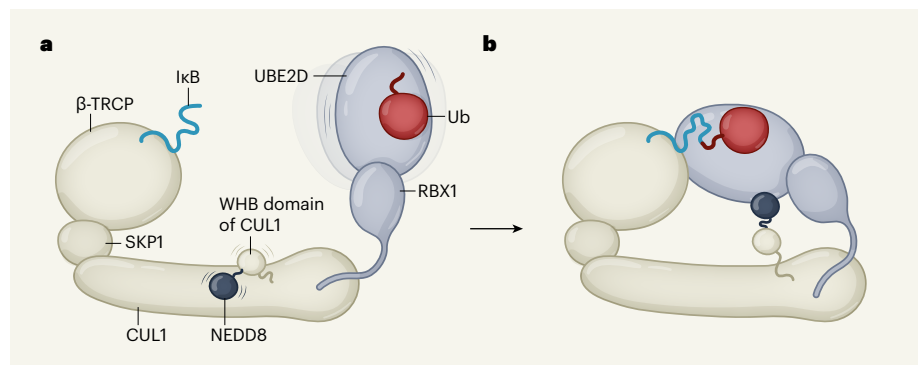


Figure 1 | Structural basis for how ubiquitination is stimulated by the NEDD8 protein. Baek *et al.*⁶ used cryo-electron microscopy to analyse how the ubiquitin protein (Ub) becomes attached to a protein that is thereby marked for degradation. **a**, Ubiquitin binds to the enzyme UBE2D. The protein IκB is a ubiquitination target, and binds to a substrate receptor called β-TRCP. This receptor also binds to a protein complex consisting of SKP1–CUL1–RBX1, called CRL1, to form a complex termed CRL1^{β-TRCP}. The transfer of ubiquitin from UBE2D to IκB is aided by CRL1^{β-TRCP}. The NEDD8 protein tags the WHB domain of CUL1, thereby increasing the flexibility of the complex and enhancing ubiquitin transfer. **b**, The authors describe a transition-state complex consisting of three modules: an activation module (the NEDD8-bound WHB domain), a catalytic module (ubiquitin, UBE2D and the adjacent part of RBX1) and a substrate-scaffolding module (the remaining components). They report that extensive rearrangements of these modules occur after NEDD8 binds to CRL1, a finding that helps to explain how NEDD8 enhances ubiquitination. (Image based on Extended Data Fig. 2 of ref. 6.)

interactions stabilize the configurations of the WHB and RING domains and bring UBE2D's active site into close proximity with β -TRCP and its bound target protein.

To confirm these findings, the authors performed extensive and sophisticated kinetic analyses comparing wild-type complexes with those containing mutant proteins designed to disrupt interactions between the modules. All complexes containing a single mutant protein showed strongly reduced enzymatic activity compared with those of wild-type complexes, and complexes containing two mutant proteins had potent synergistic defects, which is consistent with the authors' model for how the complex functions.

This structure provides information that explains many previously confusing or contradictory observations. For example, it now makes sense why, during a bacterial infection, there is a catastrophic effect on CRL function when bacterial enzymes target the glutamine 40 amino-acid residue of NEDD8 (ref. 8). This is because modification of this residue would destabilize the activation module. In addition, the structure shows clearly how direct contacts between NEDD8 and UBE2D that occur away from UBE2D's catalytic site⁹ work together with RBX1 to optimally position the catalytic module relative to the β -TRCP-bound target protein.

These structural insights pose new questions. Most notably, why does the transfer of the first ubiquitin to some CRL substrates require an extra RBX1-interacting complex of E3 and E2 enzymes (ARIH1 and UBE2L3, respectively¹⁰), given the extraordinary catalytic efficiency of the complex reported by Baek and colleagues? Moreover, how is the observed rapidity of ubiquitin transfer achieved, given the proposed requirement for the complex to undergo substantial structural rearrangements to reach the transition state? And what might the transition state look like for the NEDD8-stimulated process of chain elongation (the attachment of further ubiquitin molecules to the initial ubiquitin tag on a target protein), considering that ubiquitin-chain elongation is mediated by different E2 enzymes¹¹ from those that add the initial ubiquitin tag? With cryo-EM now firmly part of the toolkit for investigating ubiquitination, the answers might arrive sooner than we thought.

Importantly, these new structural data might help in the design of drugs known as proteolysis-targeting chimaeras (PROTACs), some of which can redirect specific CRL enzymes to ubiquitinate and thus destroy targets of clinical interest that are outside the enzymes' natural repertoire¹². These drugs work by simultaneously binding substrate receptors of CRLs and a target protein. However, the

formation of such complexes is not always sufficient to stimulate ubiquitin transfer¹³. The reason for this might become clear from the deeper understanding of CRL-mediated ubiquitin transfer gained through the work of Baek and colleagues.

Raymond J. Deshaies is at Amgen Research, Amgen, Thousand Oaks, California 91320, USA. **Nathan W. Pierce** is at Amgen Research, Amgen, South San Francisco, California 94080, USA.
e-mails: rdii2003@gmail.com; npierce@amgen.com

1. Lydeard, J. R., Schulman, B. A. & Harper, J. W. *EMBO Rep.* **14**, 1050–1061 (2013).
2. Saha, A. & Deshaies, R. J. *Mol. Cell* **32**, 21–31 (2008).
3. Duda, D. M. et al. *Cell* **134**, 995–1006 (2008).
4. Read, M. A. et al. *Mol. Cell. Biol.* **20**, 2326–2333 (2000).
5. Podust, V. N. et al. *Proc. Natl Acad. Sci. USA* **97**, 4579–4584 (2000).
6. Baek, K. et al. *Nature* **578**, 461–466 (2020).
7. Zheng, N. et al. *Nature* **416**, 703–709 (2002).
8. Cui, J. et al. *Science* **329**, 1215–1218 (2010).
9. Sakata, E. et al. *Nature Struct. Mol. Biol.* **14**, 167–168 (2007).
10. Scott, D. C. et al. *Cell* **166**, 1198–1214 (2016).
11. Sievers, Q. L., Gasser, J. A., Cowley, G. S., Fischer, E. S. & Ebert, B. L. *Blood* **132**, 1293–1303 (2018).
12. Verma, R., Mohl, D. & Deshaies, D. J. *Mol. Cell* <https://doi.org/10.1016/j.molcel.2020.01.010> (2020).
13. Smith, B. E. et al. *Nature Commun.* **10**, 131 (2019).

The authors declare competing financial interests: see go.nature.com/377wpbi for details.

This article was published online on 12 February 2020.

nature masterclasses



Free online course for peer reviewers

For researchers new to peer review or wishing to develop their skills

Register for free access at masterclasses.nature.com

Taught by Nature Research editors • 3–4 hours' learning • Free completion certificate

[W masterclasses.nature.com](https://masterclasses.nature.com)

[Follow us on LinkedIn](#)

[Skills and Careers Forum for Researchers](#)

A80873

Investigation of the fine structure of antihydrogen

<https://doi.org/10.1038/s41586-020-2006-5>

The ALPHA Collaboration*

Received: 7 December 2018

Accepted: 20 December 2019

Published online: 19 February 2020

Open access

 Check for updates

At the historic Shelter Island Conference on the Foundations of Quantum Mechanics in 1947, Willis Lamb reported an unexpected feature in the fine structure of atomic hydrogen: a separation of the $2S_{1/2}$ and $2P_{1/2}$ states¹. The observation of this separation, now known as the Lamb shift, marked an important event in the evolution of modern physics, inspiring others to develop the theory of quantum electrodynamics^{2–5}. Quantum electrodynamics also describes antimatter, but it has only recently become possible to synthesize and trap atomic antimatter to probe its structure. Mirroring the historical development of quantum atomic physics in the twentieth century, modern measurements on anti-atoms represent a unique approach for testing quantum electrodynamics and the foundational symmetries of the standard model. Here we report measurements of the fine structure in the $n = 2$ states of antihydrogen, the antimatter counterpart of the hydrogen atom. Using optical excitation of the $1S$ – $2P$ Lyman- α transitions in antihydrogen⁶, we determine their frequencies in a magnetic field of 1 tesla to a precision of 16 parts per billion. Assuming the standard Zeeman and hyperfine interactions, we infer the zero-field fine-structure splitting ($2P_{1/2}$ – $2P_{3/2}$) in antihydrogen. The resulting value is consistent with the predictions of quantum electrodynamics to a precision of 2 per cent. Using our previously measured value of the $1S$ – $2S$ transition frequency^{6,7}, we find that the classic Lamb shift in antihydrogen ($2S_{1/2}$ – $2P_{1/2}$ splitting at zero field) is consistent with theory at a level of 11 per cent. Our observations represent an important step towards precision measurements of the fine structure and the Lamb shift in the antihydrogen spectrum as tests of the charge–parity–time symmetry⁸ and towards the determination of other fundamental quantities, such as the antiproton charge radius^{9,10}, in this antimatter system.

The fine-structure splitting of the $n = 2$ states of hydrogen is the separation of the $2P_{3/2}$ and $2P_{1/2}$ levels at zero magnetic field. This splitting, predicted by the Dirac theory of relativistic quantum mechanics¹¹, originates from the spin–orbit interaction between the non-zero orbital angular momentum ($L = 1$) and the electron spin. The ‘classic’ Lamb shift is defined as the splitting between the $2S_{1/2}$ and $2P_{1/2}$ states at zero field¹², and is a manifestation of the interaction of the electron with the quantum fluctuations of the vacuum electromagnetic field, an effect explained by quantum electrodynamics (QED)^{12–14}. Today, it is understood that the classic Lamb shift in hydrogen is dominated by the QED effects on the $2S$ energy level, and that the $1S$ level receives even stronger QED corrections than the $2S$ level^{12,13}. Although QED corrections in levels $n \neq 2$ are now also sometimes referred to as Lamb shifts, in this Article we restrict our definition of the Lamb shift to be the classic $n = 2$ shift.

In a magnetic field, the Zeeman effect causes the $2P_{3/2}$ state to also split into four sublevels (labelled $2P_a$, $2P_b$, $2P_c$ and $2P_d$), whereas the $2S_{1/2}$ and $2P_{1/2}$ states each split into two ($2S_{ab}$ and $2S_{cd}$; $2P_e$ and $2P_f$). These fine-structure levels further split into two hyperfine states owing to the proton spin (see Fig. 1 for the expected energy levels for the case of antihydrogen, where the spin orientations are reversed with respect to those of hydrogen.)

Lamb’s original work used the then newly developed techniques of an excited-state atomic hydrogen beam and resonant microwave spectroscopy to study direct transitions between the $n = 2$ fine-structure states in various magnetic fields. The Lamb shift was then determined to 10% precision by extrapolating frequency measurements to zero field¹. Here, we report the observation of the splitting between the $2P_c$ and $2P_f$ states in antihydrogen in a field of 1 T, by studying laser-induced transitions from the ground state. Assuming the validity of the Zeeman and hyperfine interactions, and using the value of the previously measured $1S$ – $2S$ transition frequency⁷, we infer from our results the values of the zero-field fine-structure splitting and the classic Lamb shift in antihydrogen. Such studies have become possible owing to the combination of several recent advances: the accumulation¹⁵ of hundreds of anti-atoms in each run, their confinement for many hours¹⁶, control of the hyperfine polarization of the antihydrogen samples¹⁷ and the development of a narrow-line, pulsed, Lyman- α laser^{6,18}.

Details of the production, trapping and control of antihydrogen in the ALPHA experiment have been provided elsewhere^{6,7,15–25}, so the following description is brief. The ALPHA-2 apparatus (Fig. 2) incorporates a cylindrical magnetic trapping volume (about 400 cm³) for neutral anti-atoms; the magnetic-field minimum at the centre of the trap was

*A list of participants and their affiliations appears at the end of the paper

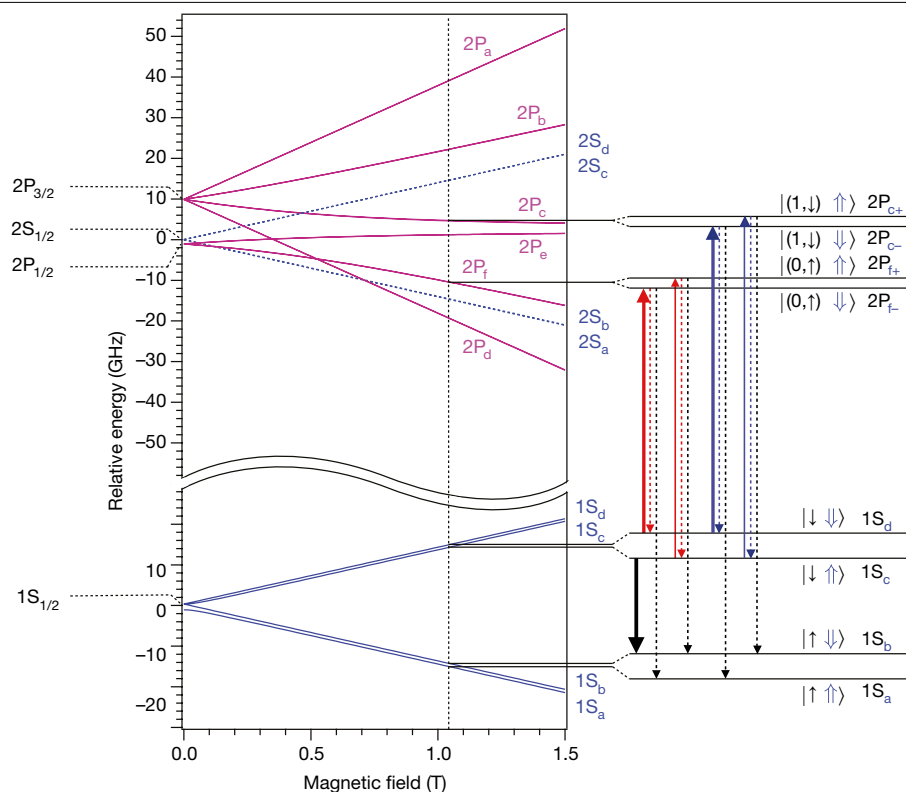


Fig. 1 | Expected antihydrogen energy levels. Calculated energies of the fine structure and the hyperfine sublevels of the $1S_{1/2}$, $2S_{1/2}$, $2P_{3/2}$ and $2P_{1/2}$ states are shown as functions of magnetic-field strength. The spin orientations for antihydrogen are shown; they are reversed for hydrogen. The centroid energy difference, $E_{1S-2S} = 2.4661 \times 10^{15}$ Hz, has been suppressed on the vertical axis. Details of the energy levels relevant to this work at a magnetic field of $B = 1.0329$ T are shown on the right. Each state is labelled using conventional notation. For the 1S and 2S states, the hyperfine states are labelled with subscripts a–d in order of increasing energy (see, for example, ref. ⁷); namely, $S_a = |\uparrow\uparrow\rangle$, $S_b = |\uparrow\downarrow\rangle$, $S_c = |\downarrow\uparrow\rangle$ and $S_d = |\downarrow\downarrow\rangle$, where the ket notation represents the positron spin (left; \downarrow or \uparrow) and antiproton spin (right; \downarrow or \uparrow) states in the high-field limit. The labels S_{ab} and S_{cd} are used when the antiproton spins are unpolarized. For the 2P states, the fine-structure splittings are labelled with

subscripts a–f in order of decreasing energy at low magnetic fields, whereas the hyperfine splitting due to the antiproton spin is specified by subscripts + and – for spin parallel ($\uparrow\uparrow$) and anti-parallel ($\downarrow\downarrow$) to the magnetic field in the high-field limit, respectively. The symbol ($\downarrow\uparrow$) in the figure indicates that the positron spin states are mixed for the $2P_c$ and $2P_f$ states. The vertical solid arrows indicate the one-photon laser transitions probed here: $1S_d \rightarrow 2P_f$ (bold red), $1S_c \rightarrow 2P_f$ (thin red), $1S_d \rightarrow 2P_{c-}$ (bold blue) and $1S_c \rightarrow 2P_{c+}$ (thin blue). The dashed red and blue arrows indicate relaxation to the same trappable level, which is not detectable in the present experiment, and the dashed black arrows indicate relaxation to untrappable levels, which is detectable via annihilation signals (see text). The bold black arrow shows the microwave transition used to eliminate $1S_c$ state atoms to prepare a doubly spin-polarized antihydrogen sample.

set to 1.0329 ± 0.0004 T for this work. (All uncertainties given herein are 1σ .) By combining 90,000 trapped antiprotons from the CERN Antiproton Decelerator²³ and three million positrons from a positron accumulator^{24,25}, about 10–30 cold (below 0.54 K) anti-atoms are confined in the magnetic trap in a 4-min cycle. Under normal conditions, the storage lifetime¹⁶ of the trapped antihydrogen is greater than 60 h, which permits loading from repeated cycles¹⁵ to obtain hundreds of antihydrogen atoms in a few hours.

Two types of antihydrogen samples were used in these studies. The positron spin of an antihydrogen atom confined in the ALPHA-2 trap is necessarily polarized, because only the $1S_c$ and $1S_d$ states can be magnetically trapped (Fig. 1). The antiproton spin, on the other hand, is unpolarized a priori, with both orientations equally likely. Thus, the initial samples are singly spin-polarized. On the other hand, doubly spin-polarized samples, in which both the positron and antiproton spins are polarized, can be prepared by injecting microwaves to resonantly drive the $1S_c$ atoms to the untrappable $1S_b$ state (Fig. 1), effectively depopulating the $1S_c$ state from the trap¹⁷.

Spectroscopy in the vacuum ultraviolet range is challenging even for ordinary atoms, owing in part to the lack of convenient laser sources and optical components^{26–28}. Our pulsed, coherent 121.6-nm radiation was produced by generating the third harmonic of 365-nm pulses in a Kr/Ar gas mixture at a repetition rate of 10 Hz (ref. ¹⁸). The typical pulse width

at 121.6 nm was 12 ns, and the bandwidth was estimated from the Fourier transform of the temporal pulse shape to be 65 MHz (full-width at half-maximum, FWHM). The 121.6-nm light was linearly polarized because of the three-photon mixing of linearly polarized 365-nm light. In the antihydrogen trap, the polarization vector was nearly perpendicular to the direction of the axial magnetic field. The laser beam had a radius of 3.6 mm and was roughly collimated across the trapping region (Fig. 2). The average pulse energies in the antihydrogen trapping volume ranged from 0.44 nJ to 0.72 nJ over different runs, as evaluated from the pulse waveforms recorded with a calibrated, solar-blind photomultiplier detector.

In this experiment, single-photon transitions from the $1S_c$ ($1S_d$) states to the $2P_{c+}$ ($2P_{c-}$) and $2P_{f+}$ ($2P_{f-}$) states are driven by the 121.6-nm light (red and blue arrows in Fig. 1). When antihydrogen is excited to the $2P_{c+}$ or $2P_{f+}$ state, it decays to the ground-state manifold within a few nanoseconds by emitting a photon at 121.6 nm. The mixed nature of the positron spin states in the $2P_{c+}$ ($2P_{c-}$) and $2P_{f+}$ ($2P_{f-}$) states implies that these states can decay to the $1S_b$ ($1S_a$) states via a positron spin flip (black dashed arrows in Fig. 1). Atoms in these final states are expelled from the trap and are annihilated on the trap walls. Annihilation products (charged pions) are in turn detected by a silicon vertex detector²⁹ with an efficiency greater than 80%.

Table 1 summarizes our data. In total, four series of measurements were performed using either singly or doubly spin-polarized samples.

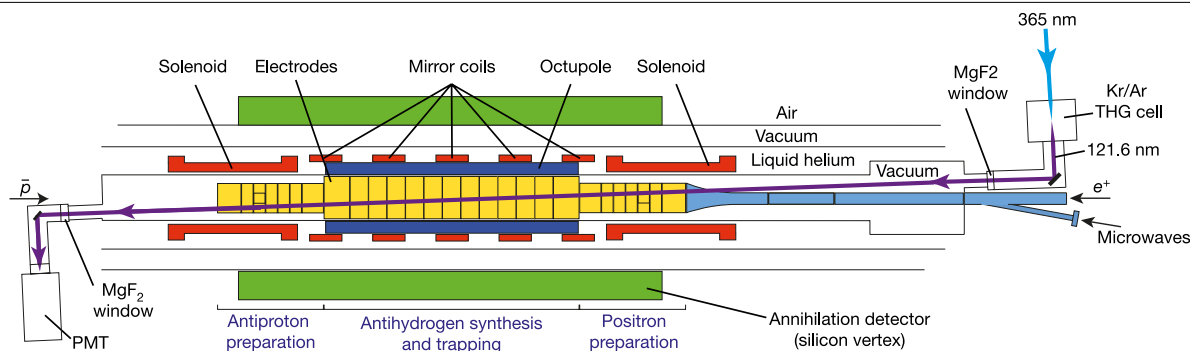


Fig. 2 | The ALPHA-2 central apparatus. A cylindrical trapping volume for neutral antimatter with a diameter of 44.35 mm and an axial length of 280 mm is located inside several Penning trap electrodes and surrounded by an octupole coil, five mirror coils and two solenoids, all superconducting. The three-layer silicon vertex annihilation detector is shown schematically in green. Laser light (purple line) enters from the positron (e^+) side (right) and is transmitted to the antiproton (\bar{p}) side (left) through vacuum-ultraviolet-grade

MgF₂ ultrahigh-vacuum windows. The laser beam crosses the trap axis at an angle of 2.3°. The transmitted 121.6-nm pulses are detected by a solar-blind photomultiplier tube (PMT) at the antiproton side. Microwaves used to prepare the doubly spin-polarized samples are introduced from the positron side through a waveguide, shown in blue. The external solenoid magnet for the Penning traps is not shown here. THG, third-harmonic generation.

The Series 1 data, previously reported in ref. ⁶, have been reanalysed. Each series consisted of two or four runs, and in each run about 500 antihydrogen atoms were accumulated over approximately two hours, typically involving over 30 production cycles. The trapped anti-atoms were then irradiated for about two hours by a total of 72,000 laser pulses at twelve different frequencies (that is, 6,000 pulses per frequency point for each run) spanning the range -3.10 GHz to $+2.12$ GHz relative to the expected (hydrogen) transition frequencies. The laser frequency was changed every 20 s in a non-monotonic fashion to minimize effects related to the depletion of the sample of antihydrogen. After the laser exposure, the remaining antihydrogen atoms were released by shutting down the trap magnets, typically in 15 s, and counted via detection of their annihilation events. 40–60% of the trapped antihydrogen atoms experienced resonant, laser-induced spin flips, and their annihilations were detected during the two-hour laser irradiation period.

A combination of time-gated antihydrogen detection (enabled by the use of a pulsed laser), the accumulation of a large number of anti-atoms and the use of supervised machine-learning analysis²⁹ (based on a boosted decision-tree classifier) suppressed the background to a negligible level (less than 2 counts per 2-h irradiation period).

The measured spectra, obtained from counting the laser-induced spin-flip events, are shown in Fig. 3 for both singly and doubly spin-polarized antihydrogen samples. For each run, the probability at each frequency point is determined from dividing the number of annihilation events recorded at that frequency by the total number of trapped atoms in that run, and further dividing by the ratio of the average laser energy to a standard value of 0.5 nJ. The normalization to the standard laser energy is to account for the expected linear dependence of the transition probability on the laser power in our regime. The data plotted in Fig. 3 are spectrum-averaged over the runs for each series. For the singly polarized sample (Fig. 3a), each transition shows a linewidth of

about 1.5 GHz (FWHM). This is consistent with the expected Doppler broadening in our trapping condition (1 GHz FWHM) and the hyperfine splitting of the $1S-2P_f$ and $1S-2P_c$ transitions (0.71 GHz for both transitions). The hyperfine structure cannot be resolved in these singly polarized samples owing to the Doppler broadening.

Figure 3b shows the spectra obtained from doubly spin-polarized antihydrogen samples. For these data, microwave radiation of -28 GHz (power ~ 0.4 W, measured at the trap entrance) was applied before the start of optical spectroscopy, in the form of a 9-MHz sweep, covering the $1S_c-1S_b$ transition in the magnetic-field minimum¹⁷. As shown in Table 1, about half of the total trapped antihydrogen atoms underwent a positron spin-flip and annihilated during microwave irradiation. This is consistent with our experience from earlier studies, in which $1S_c$ -state atoms were removed with about 95% efficiency^{7,17}. The spectral lines of the $1S-2P$ transitions in doubly spin-polarized antihydrogen (Fig. 3b) are narrower than those in the singly spin-polarized samples (Fig. 3a) because the former involves only one hyperfine state in the ground state. The peaks are red-shifted because the frequencies of the transition from the $1S_d$ state to the $2P_f$ and $2P_c$ states are expected to be about 700 MHz lower than those from the $1S_c$ state. The observed width of ~ 1 GHz FWHM of these lines is in agreement with the Doppler width expected for our trapping conditions.

The procedure used to extract the frequencies of the fine-structure transitions and to evaluate their associated uncertainties is described in Methods. We summarize the results of this analysis in Table 2. A simulation was used to model the motion of trapped antihydrogen atoms in the ALPHA-2 trap and their interaction with pulsed laser radiation. The resonance transition frequencies were obtained by comparing simulated and experimental lineshapes. Extensive investigations were performed to evaluate systematic uncertainties in our measurement (Table 3). The validity of our analysis procedure was tested by using

Table 1 | Experimental parameters and number of detected events

Series	Sample polarization	Transition probed	Number of runs	Average pulse energy (pJ)	Number of frequencies	Number of pulses per frequency	Number of trapped atoms	Microwave counts	Laser counts	Counts upon release
1	Single	$1S_{cd} \rightarrow 2P_{ct}$	4	600	12	24,000	2,004	–	1,197	807
2	Single	$1S_{cd} \rightarrow 2P_{ft}$	4	550	12	24,000	2,012	–	1,075	937
3	Double	$1S_d \rightarrow 2P_{c-}$	2	440	12	12,000	1,044	527	229	288
4	Double	$1S_d \rightarrow 2P_{f-}$	2	720	12	12,000	971	463	341	167

The experimental parameters, together with the number of antihydrogen events detected during the microwave irradiation, the laser irradiation and the release of the remaining atoms, are tabulated for each series. The machine-learning analysis identifies annihilation events with an estimated efficiency of 0.849 for the microwave irradiation, 0.807 for the laser irradiation and 0.851 for the release of the remaining atoms. The number of counts is corrected for the detection efficiencies. The number of trapped atoms is derived from the sum of the other counts.

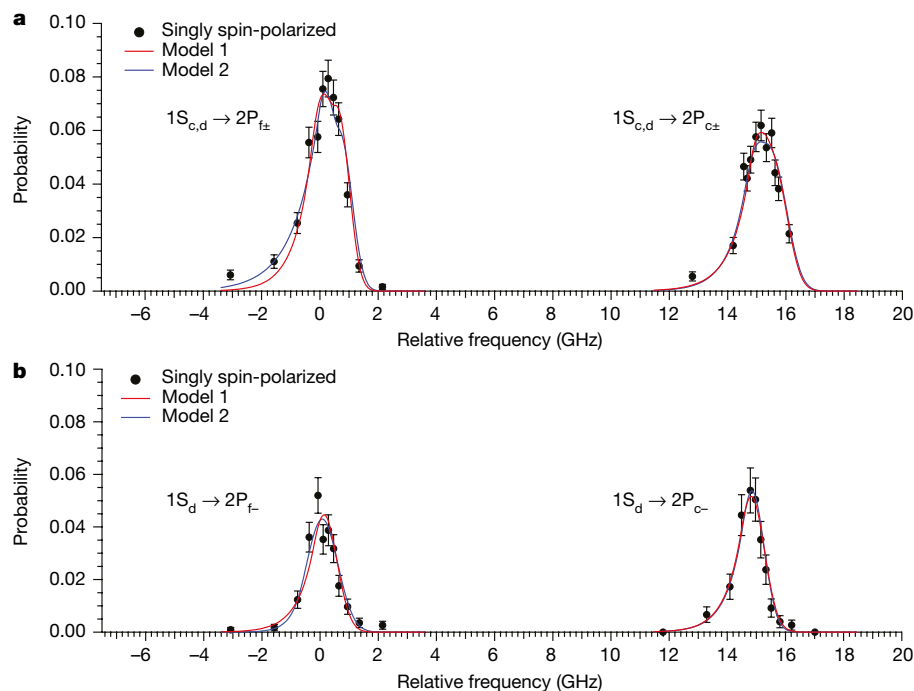


Fig. 3 | 1S–2P fine-structure spectrum of antihydrogen. **a, b**, Experimental data (filled circles) and fitted lineshapes for singly spin-polarized (**a**) and doubly spin-polarized (**b**) antihydrogen samples. The data points were obtained from the detected spin-flip events, normalized to the total number of trapped antihydrogen atoms, for a laser pulse energy of 0.5 nJ. The error bars are 1σ counting uncertainties. The frequency is offset by 2,466,036.3 GHz. We

note that no data were taken between the two peaks (~ 2 – 12 GHz). The red fit curves were obtained via our standard fitting procedure (Model 1), and the blue curves were derived from an alternative fitting model (Model 2), illustrating the sensitivity of our results to the fitting procedure. See text and Methods for detailed discussion.

different lineshape-fitting models. Two representative curve fits are shown in Fig. 3. The fit of Model 1 uses a function constrained to fit the simulation shape, whereas in Model 2 the shape parameters of this function are allowed to vary to best fit the experimental data; see Methods for details. The sensitivity of the results to the experimental and simulation parameters was tested by repeating the analysis procedure for a number of simulations with varied input. These included the initial antihydrogen conditions (such as the initial temperature, the quantum state, and the cloud diameter of antihydrogen at formation) and laser properties (such as linewidth, beam waist size and beam position); see Methods and Extended Data Fig. 1. Other sources of systematic uncertainties include the calibration accuracy and a possible frequency drift of the wavemeter, frequency shifts of the 730-nm amplification laser cavity, and possible incomplete clearing of the $1S_c$ state in the preparation of the doubly spin-polarized samples (Table 3 and Methods).

Within the uncertainties, the measured transition frequencies agree with theoretical expectations for hydrogen for all four series (Table 2, Fig. 4). The fact that the four measurements are consistent, despite having different systematics, increases the confidence in our overall results. The results can be combined to give a test of charge–parity–time (CPT)

invariance in the 1S–2P transitions at the level of 16 parts per billion (Fig. 4).

Fundamental physical quantities of antihydrogen can be extracted from our optical measurements of the 1S–2P transitions by combining them with our earlier measurement of the 1S–2S transition in the same magnetic trapping field⁷. From the weighted average of the results between the singly polarized and doubly polarized measurements (Table 1), we obtain a $2P_{c-}-2P_{f-}$ splitting of 14.945 ± 0.075 GHz, a $2S_d-2P_{c-}$ splitting of 9.832 ± 0.049 GHz and a $2S_d-2P_{f-}$ splitting of 24.778 ± 0.060 GHz at 1.0329 T (Methods). Only two of these three splittings are independent, and they all agree with the values predicted for hydrogen in the same field.

In interpreting our data, we categorize features in the spectrum based on the order of the fine-structure constant α in a perturbative series expansion in quantum field theory (which is assumed to be valid for the purpose of our categorization). Those features that can be described by the Dirac theory (the Zeeman, hyperfine and fine-structure effects) are referred to as ‘tree-level effects’ and follow from the lower-order terms (up to order $\sim \alpha^2 Ry$, where Ry is the Rydberg constant). On the other hand, the Lamb shift originates from the so-called ‘loop effects’ (order $\sim \alpha^3 Ry$), the calculation of which requires the concept of renormalization

Table 2 | 1S–2P transition frequencies

	Sample spin polarization	Antihydrogen $f_{\text{res}}(\text{exp})$ (MHz)	Hydrogen $f_{\text{res}}(\text{th})$ (MHz)	Difference $f_{\text{res}}(\text{exp}) - f_{\text{res}}(\text{th})$ (MHz)
$1S_{c,d} \rightarrow 2P_{c\pm}$	Single	2,466,051,659(62)	2,466,051,625	34
$1S_{c,d} \rightarrow 2P_{f\pm}$	Single	2,466,036,611(88)	2,466,036,642	–31
$1S_d \rightarrow 2P_{c-}$	Double	2,466,051,189(76)	2,466,051,270	–81
$1S_d \rightarrow 2P_{f-}$	Double	2,466,036,395(81)	2,466,036,287	108

The experimentally determined transition frequencies for antihydrogen $f_{\text{res}}(\text{exp})$ (with 1σ errors in parentheses) are compared with the theoretically expected values for hydrogen $f_{\text{res}}(\text{th})$ at a magnetic field of 1.0329 T. For the singly spin-polarized data, the centroid of the hyperfine states is given. The transition frequencies for hydrogen were calculated to a precision better than 1 MHz (Methods).

Table 3 | Summary of uncertainties

Source of uncertainty	$1S_d \rightarrow 2P_{c-}$ Doubly spin-polarized (MHz)	$1S_d \rightarrow 2P_{f-}$ Doubly spin-polarized (MHz)	$1S_{cd} \rightarrow 2P_{c\pm}$ Singly spin-polarized (MHz)	$1S_{cd} \rightarrow 2P_{f\pm}$ Singly spin-polarized (MHz)
Lineshape fit statistics	55	54	45	47
Fitting-model dependence	24	42	17	62
Wavemeter drift	30	30	30	30
Wavemeter offset	18	18	18	18
730-nm cavity frequency correction	18	18	18	18
Residual $1S_c$ state atoms in doubly spin-polarized sample	23	16	0	0
Magnetic field	5	8	5	8
Total	76	81	62	88

Estimated uncertainties (1 σ) at 121.6 nm for each transition (Methods).

to avoid infinities^{12–14}. Each of the measured splittings has different sensitivity to different terms. At the level of our precision, the $2P_{c-} - 2P_{f-}$ splitting is sensitive to the tree-level terms with negligible QED effects, whereas the $2S - 2P_f$ and $2S - 2P_c$ splittings are sensitive to the field-independent Lamb shift, in addition to the tree-level terms (we note that the Lamb shift is predicted to have negligible dependence on the magnetic field¹⁴). The agreement between our measurement and the Dirac prediction for the $2P_{c-} - 2P_{f-}$ splitting supports the consistency of the tree-level theory in describing the Zeeman, hyperfine and fine-structure interactions in the 2P states of antihydrogen. If we hence assume that we can correctly account for the tree-level effects in our measurements, we can infer from our measured splittings the values of the zero-field fine-structure splitting in antihydrogen to be 10.88 ± 0.19 GHz. By combining the current result with the much more precisely measured 1S–2S transition frequency in antihydrogen⁷, we obtain a classic Lamb shift of 0.99 ± 0.11 GHz (Methods). If we use the theoretical value of the fine-structure splitting from the Dirac prediction (rather than treat it as a parameter), we can derive a tighter constraint on the Lamb shift, 1.046 ± 0.035 GHz.

When considering the first measurements on an exotic system such as antihydrogen, it is necessary to adopt a framework within which it is possible to compare the results to the expectations of well established models for normal matter. The choice of which effects can be assumed to be true in interpreting the data are, of necessity, somewhat arbitrary. The approach illustrated here is based on the order of perturbation in the coupling constant α ; we have assumed (lower-order) tree-level effects in order to extract (higher-order) renormalizable loop effects.

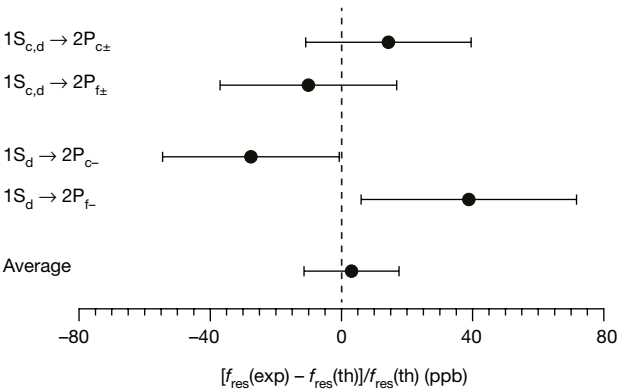


Fig. 4 | Comparison of antihydrogen and hydrogen transition frequencies. The experimentally measured frequencies for the 1S–2P transitions in antihydrogen $f_{\text{res}}(\text{exp})$ are compared with those theoretically expected for hydrogen $f_{\text{res}}(\text{th})$ (Table 2). All four measurements are consistent with hydrogen, and their average gives a combined test of CPT invariance at 16 parts per billion (ppb). The error bars are 1σ , and the calculation of the error bar for the average takes into account correlated uncertainties (Methods).

Other approaches are possible in interpreting our data. We note that if the standard theory for the hydrogen atom applies to antihydrogen, most of the expected QED effect is on the 2S level, rather than on the 2P level. Furthermore, the 1S level receives approximately $n^3 = 8$ times larger QED corrections than the 2S level; hence, our earlier accurate determination of the antihydrogen 1S–2S level difference⁷ gives strong constraints on new interactions within the QED framework. However, it is possible that a new effect could show up in the antihydrogen classic Lamb shift while satisfying the 1S–2S constraint. See ref.⁸ for an example in a Lorentz-violating effective-field theory framework.

We have investigated the fine structure of the antihydrogen atom in the $n = 2$ states. The splitting between the $2P_c$ and $2P_f$ states, two of the 2P Zeeman sublevels belonging to the $J = 3/2$ and $J = 1/2$ manifolds (J , total angular momentum), has been observed in a magnetic field of 1 T. The energy levels of the 1S–2P transitions agree with the Dirac theory predictions for hydrogen at 1 T to 16 parts per billion, and their difference to 0.5%. By assuming the standard Zeeman and hyperfine effects, and by combining our results with the earlier result of 1S–2S spectroscopy⁷, we have inferred the zero-field fine-structure splitting and the classic Lamb shift in the $n = 2$ level.

These observations expand the horizons of antihydrogen studies, providing opportunities for precision measurements of the fine structure and the Lamb shift—both of which are longstanding goals in the field. Prospects exist for considerable improvements in the precision beyond this initial determination. With the advent of the ELENA ring in 2021, an upgrade to the Antiproton Decelerator with an anticipated increase in the antiproton flux, the statistical uncertainties are expected to be dramatically reduced. The development of laser cooling³⁰ would reduce the Doppler width to a level comparable to the natural linewidth, which in turn would improve the precision of the frequency determination. It would also permit direct experimental determination of the hyperfine splitting in the 2P states, for which theoretical values were assumed in this study.

Such measurements will provide tests of CPT invariance that are complementary to other precision measurements in antihydrogen, such as the 1S–2S frequency and the ground-state hyperfine splitting. Furthermore, a precise value of the classic Lamb shift, combined with that of the 1S–2S interval, will permit an antimatter-only determination of the antiproton charge radius^{9,10}, without referring to matter measurements—that is, independent of the proton charge radius puzzle^{31–33}. These examples signify the importance of broad and complementary measurements in testing fundamental symmetries. In the absence of compelling theoretical arguments to guide the way to possible asymmetries, it is essential to address the antihydrogen spectrum as comprehensively as is practical. Finally, the results reported here demonstrate our capability to precisely and reproducibly drive vacuum ultraviolet transitions on a few anti-atoms, and indicate our readiness for laser cooling of antihydrogen³⁰, an eagerly anticipated development in antimatter studies with far-reaching implications for both spectroscopic and gravitational studies³⁴.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2006-5>.

- Lamb, W. E., Jr & Retherford, R. C. Fine structure of the hydrogen atom by a microwave method. *Phys. Rev.* **72**, 241–243 (1947).
- Tomonaga, S. On a relativistically invariant formulation of the quantum theory of wave fields. *Prog. Theor. Phys.* **1**, 27–42 (1946).
- Schwinger, J. On quantum-electrodynamics and the magnetic moment of the electron. *Phys. Rev.* **73**, 416–417 (1948).
- Feynman, R. P. Space-time approach to quantum electrodynamics. *Phys. Rev.* **76**, 769–789 (1949).
- Schweber, S. S. *QED and the Men who Made it: Dyson, Feynman, Schwinger, and Tomonaga* (Princeton Univ. Press, 1994).
- Ahmadi, M. et al. Observation of the 1S–2P Lyman- α transition in antihydrogen. *Nature* **561**, 211–215 (2018).
- Ahmadi, M. et al. Characterization of the 1S–2S transition in antihydrogen. *Nature* **557**, 71–75 (2018).
- Kostecký, V. A. & Vargas, A. J. Lorentz and CPT tests with hydrogen, antihydrogen, and related systems. *Phys. Rev. D* **92**, 056002 (2015).
- Crivelli, P., Cooke, D. & Heiss, M. W. Antiproton charge radius. *Phys. Rev. D* **94**, 052008 (2016).
- Eriksson, S. Precision measurements on trapped antihydrogen in the ALPHA experiment. *Philos. Trans. Royal Soc. A* **376**, 20170268 (2018).
- Dirac, P. A. M. The quantum theory of the electron. *Proc. R. Soc. A* **117**, 610–624 (1928).
- Kinoshita, T. *Quantum Electrodynamics* (World Scientific, 1990).
- Karshenboim, S. G. Precision physics of simple atoms: QED tests, nuclear structure and fundamental constants. *Phys. Rep.* **422**, 1–63 (2005).
- Brodsky, S. J. & Parsons, R. G. Precise theory of the Zeeman spectrum for atomic hydrogen and deuterium and the Lamb shift. *Phys. Rev.* **163**, 134–146 (1967).
- Ahmadi, M. et al. Antihydrogen accumulation for fundamental symmetry tests. *Nat. Commun.* **8**, 681 (2017).
- Capra, A. & ALPHA Collaboration. Lifetime of magnetically trapped antihydrogen in ALPHA. *Hyperfine Interact.* **240**, 9 (2019).
- Ahmadi, M. et al. Observation of the hyperfine spectrum of antihydrogen. *Nature* **548**, 66–69 (2017); erratum **553**, 530 (2018).
- Michan, J. M., Poloviy, G., Madison, K. W., Fujiwara, M. C. & Momose, T. Narrowband solid state VUV coherent source for laser cooling of antihydrogen. *Hyperfine Interact.* **235**, 29–36 (2015).
- Andresen, G. B. et al. Trapped antihydrogen. *Nature* **468**, 673–676 (2010).
- Andresen, G. B. et al. Confinement of antihydrogen for 1,000 seconds. *Nat. Phys.* **7**, 558–564 (2011).
- Ahmadi, M. et al. Observation of the 1S–2S transition in trapped antihydrogen. *Nature* **541**, 506–510 (2017).
- Ahmadi, M. et al. Enhanced control and reproducibility of non-neutral plasmas. *Phys. Rev. Lett.* **120**, 025001 (2018).
- Maury, S. The antiproton decelerator: AD. *Hyperfine Interact.* **109**, 43–52 (1997).
- Murphy, T. J. & Surko, C. M. Positron trapping in an electrostatic well by inelastic collisions with nitrogen molecules. *Phys. Rev. A* **46**, 5696–5705 (1992).
- Surko, C. M., Greaves, R. G. & Charlton, M. Stored positrons for antihydrogen production. *Hyperfine Interact.* **109**, 181–188 (1997).
- Luiten, O. J. et al. Lyman- α spectroscopy of magnetically trapped atomic hydrogen. *Phys. Rev. Lett.* **70**, 544–547 (1993).
- Eikema, K. S. E., Walz, J. & Hänsch, T. W. Continuous coherent Lyman- α excitation of atomic hydrogen. *Phys. Rev. Lett.* **86**, 5679–5682 (2001).
- Gabrieelse, G. et al. Lyman- α source for laser cooling antihydrogen. *Opt. Lett.* **43**, 2905–2908 (2018).

- Stracka, S. Real-time detection of antihydrogen annihilations and applications to spectroscopy. *EPJ Web Conf.* **71**, 00126 (2014).
- Donnan, P. H., Fujiwara, M. C. & Robicheaux, F. A proposal for laser cooling antihydrogen atoms. *J. Phys. B* **46**, 025302 (2013).
- Pohl, R. et al. The size of the proton. *Nature* **466**, 213–216 (2010).
- Beyer, A. et al. The Rydberg constant and proton size from atomic hydrogen. *Science* **358**, 79–85 (2017).
- Fleurbay, H. et al. New measurement of the transition frequency of hydrogen: contribution to the proton charge radius puzzle. *Phys. Rev. Lett.* **120**, 183001 (2018).
- The ALPHA Collaboration & Charman, A. E. Description and first application of a new technique to measure the gravitational mass of antihydrogen. *Nat. Commun.* **4**, 1785 (2013).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

The ALPHA Collaboration

M. Ahmadi¹, B. X. R. Alves², C. J. Baker³, W. Bertsche^{4,5}, A. Capra⁶, C. Carruth⁷, C. L. Cesar⁸, M. Charlton³, S. Cohen⁹, R. Collister⁶, S. Eriksson³, A. Evans¹⁰, N. Evetts¹¹, J. Fajans⁷, T. Friesen^{2,10}, M. C. Fujiwara^{6,13}, D. R. Gill⁶, P. Granum², J. S. Hangst^{2,13}, W. N. Hardy¹¹, M. E. Hayden¹², E. D. Hunter⁷, C. A. Isaac³, M. A. Johnson^{4,5}, J. M. Jones³, S. A. Jones^{2,3}, S. Jones¹³, A. Khramov^{6,11}, P. Knapp³, L. Kurchaninov⁶, N. Madsen³, D. Maxwell³, J. T. K. McKenna^{2,5}, S. Menary¹⁴, J. M. Michan^{6,15}, T. Momose^{11,16}, J. J. Munich¹², K. Olchanski⁶, A. Olin^{6,17}, P. Pusa¹, C. Ø. Rasmussen², F. Robicheaux¹⁸, R. L. Sacramento⁸, M. Sameed⁴, E. Sarid¹⁹, D. M. Silveira⁸, C. So^{6,10}, D. M. Starko¹⁴, G. Stutter², T. D. Tharp^{19,20}, R. I. Thompson^{6,10}, D. P. van der Werf^{2,21} & J. S. Wurtele⁷

¹Department of Physics, University of Liverpool, Liverpool, UK. ²Department of Physics and Astronomy, Aarhus University, Aarhus, Denmark. ³Department of Physics, College of Science, Swansea University, Swansea, UK. ⁴School of Physics and Astronomy, University of Manchester, Manchester, UK. ⁵Cockcroft Institute, Warrington, UK. ⁶TRIUMF, Vancouver, British Columbia, Canada. ⁷Department of Physics, University of California at Berkeley, Berkeley, CA, USA. ⁸Instituto de Física, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil. ⁹Department of Physics, Ben-Gurion University of the Negev, Beer-Sheva, Israel. ¹⁰Department of Physics and Astronomy, University of Calgary, Calgary, Alberta, Canada. ¹¹Department of Physics and Astronomy, University of British Columbia, Vancouver, British Columbia, Canada. ¹²Department of Physics, Simon Fraser University, Burnaby, British Columbia, Canada. ¹³Department of Physics, Stockholm University, Stockholm, Sweden. ¹⁴Department of Physics and Astronomy, York University, Toronto, Ontario, Canada. ¹⁵École Polytechnique Fédérale de Lausanne (EPFL), Swiss Plasma Center (SPC), Lausanne, Switzerland. ¹⁶Department of Chemistry, University of British Columbia, Vancouver, British Columbia, Canada. ¹⁷Department of Physics and Astronomy, University of Victoria, Victoria, British Columbia, Canada. ¹⁸Department of Physics and Astronomy, Purdue University, West Lafayette, IN, USA. ¹⁹Soreq NRC, Yavne, Israel. ²⁰Physics Department, Marquette University, Milwaukee, WI, USA. ²¹IRFU, CEA/Saclay, Gif-sur-Yvette, France. [✉]e-mail: Makoto.Fujiwara@triumf.ca; jeffrey.hangst@cern.ch; momose@chem.ubc.ca

Methods

Transition-frequency determination

The observed 1S–2P transition spectra have asymmetric shapes with a low-frequency tail caused by Zeeman shifts in the inhomogeneous-magnetic-field regions away from the centre of the ALPHA-2 trap. As a result, the apparent peak of the observed spectrum is shifted to a slightly lower frequency with respect to the resonance transition frequency f_{res} , which is defined for atoms in resonance at the magnetic-field minimum of the trap. This offset is relatively small (of the order of 50 MHz). Nonetheless, we performed extensive analysis to understand the effects of this asymmetry on our transition-frequency determination. The details of the analysis follow.

A detailed simulation was used to model the motion of trapped antihydrogen atoms in the ALPHA-2 trap, as well as their interaction with pulsed laser radiation. Aspects of our simulation have been validated in previous studies (for example, refs. ^{10,11,19–24}). To determine the resonance transition frequency, we first simulated lineshapes for the transitions from the two trappable 1S hyperfine states to the $2P_c$ and $2P_f$ excited states (that is, for four transitions: $1S_c \rightarrow 2P_{c+}$, $1S_c \rightarrow 2P_{f+}$, $1S_d \rightarrow 2P_{c-}$ and $1S_d \rightarrow 2P_{f-}$). We then fitted each component with an asymmetric lineshape function, referred to as GE. GE is a Gaussian spliced to an exponential low frequency tail, where the derivative of the crossover point is required to be continuous. GE has four parameters: the peak frequency (f_{peak}) and the width (W) of the Gaussian, the crossover point frequency (f_x) and the overall amplitude (A). From the fit, we determined the simulated lineshape parameters $f_{\text{peak}}(\text{sim})$, $W(\text{sim})$, $f_x(\text{sim})$ and $A(\text{sim})$ for each transition. In addition, we derived the peak frequency offset Δf , defined as $\Delta f = f_{\text{peak}}(\text{sim}) - f_{\text{res}}(\text{th})$, where $f_{\text{res}}(\text{th})$ is the expected theoretical resonance frequency for hydrogen in the magnetic field B .

The experimentally observed spectra were then fitted with GE lineshapes. A sum of two GEs was used to fit singly spin-polarized samples, where only f_{peak} and a single normalization factor were used as the fitting parameters, whereas the rest of the parameters (that is, the W and f_x values of each GE, the spacing of f_{peak} between two GEs, and the ratio of the amplitudes A of two GEs) were fixed to the corresponding simulated values. For doubly spin-polarized samples, the experimental spectra were fitted with a single GE lineshape. In these fits, W and f_x were fixed using a fit to the simulated spectrum in which an estimated 5% contamination of the $1S_c$ component was assumed. The experimental transition frequency is given by $f_{\text{res}}(\text{exp}) = f_{\text{peak}}(\text{exp}) - \Delta f$, where $f_{\text{peak}}(\text{exp})$ is the peak frequency of the experimental data obtained by the fit. Here Δf corrects for the asymmetric lineshape as described earlier. The red lines (labelled as ‘Model 1’) in Fig. 3 show the results of these fits using standard simulations. We note that the transition to the $2P_c$ state is allowed when the laser polarization is not perfectly perpendicular to the B field. This could arise from the slight angle between the laser and the magnetic field (maximum 4° at the edge of our trap) or from a possible nonlinear component in the polarization of the 121-nm light (expected to be of the order of 10% or less). The frequency of the 1S– $2P_c$ transition is well separated from that of the 1S– $2P_f$ transition (by about -3.5 GHz), and its predicted intensity is very small (less than a few per cent of that for the 1S– $2P_f$ transition), hence it was ignored in the analysis.

Transition-frequency uncertainties

Extensive studies were performed to quantify the uncertainties in our frequency determination. The standard simulated spectra reproduce the observed lineshape reasonably well without any fine-tuning (Extended Data Fig. 1). The sensitivity of the obtained resonance frequency $f_{\text{res}}(\text{exp})$ to the input parameters in the simulation was studied by varying these input parameters and repeating the same analysis.

The standard input to the simulation and the range of the parameters studied (given in parentheses), were as follows. Laser pulse energies, 500 pJ (350 pJ, 800 pJ); laser line linewidth, 65 MHz (50 MHz, 80 MHz); relative magnitude of the laser sideband (present at +90 MHz with

respect to the main band owing to multimode lasing in the 730-nm amplification cavity), 10% (0%, 25%); radial position displacement of the laser beam: 0 mm (0 mm, 3 mm); initial quantum state of antihydrogen at formation: $n = 30$ (1, 30); initial diameter of the cloud of antihydrogen: 0.45 mm (0.45 mm, 0.90 mm); temperature of antihydrogen at formation (before trapping): 15 K (1 K, 15 K).

An alternative fitting method was also used to study the robustness of our procedure. Here, the lineshape function GE was fitted to the data without using constraints from the simulated spectrum. From the fit, $f_{\text{peak}}(\text{exp})$ was extracted for each transition, and the experimental resonance frequency was determined as $f_{\text{res}}(\text{exp}) = f_{\text{peak}}(\text{exp}) - \Delta f$, where the offset Δf from the standard simulation was assumed. The lineshapes given by these fits are shown by blue lines (labelled as ‘Model 2’) in Fig. 3.

The results of the analyses using the simulations with varied input parameters, as well as alternative fitting models, are given by red lines in Extended Data Fig. 1, which illustrates that the dependence on the details of the fitting procedure is small. The variations of the extracted frequency $f_{\text{res}}(\text{exp})$ in these studies (both with different simulation inputs and different fitting methods) were generally within the statistical uncertainties of these fits. We took the largest deviations in $f_{\text{res}}(\text{exp})$ among these studies as a measure of the fitting-model dependence (Table 3).

It should be noted that our evaluation of the fitting-model dependence systematics relies on the GE model being a reasonable representation of the simulated data. This agreement is qualitatively illustrated in Extended Data Fig. 1. Quantitatively, for the simulations with the standard input parameters, the χ^2 per degree of freedom (DOF) ranges from 1.2 to 2.5 (with an average of 1.8), where DOF = 8. When the input parameters are varied in the fits to the data, the χ^2 per DOF ranges from 1.0 to 3.9, with an average of 2.1. The simulation statistics were roughly a factor of 2–4 greater than the data; hence, the uncertainties arising from our analytical model of the simulation lineshape are small.

The sources of uncertainty in the transition frequencies can be summarized as follows (we note that the frequency uncertainties at 730 nm should be multiplied by a factor of 6 to give those at 121 nm): (a) Wavemeter drift: this is due to temperature-induced drift of the wavemeter readings, which was estimated from offline studies to be about 20 MHz K^{-1} at 730 nm. Given the recorded temperature variation of $\pm 0.25 \text{ K}$, we assigned an error of $\pm 5 \text{ MHz}$ at 730 nm. We note that a temperature drift during our 2-h measurements would result in a broadening of the observed linewidth. This effect would be also taken into account partly by the fitting-model uncertainty discussed above. Therefore, there is a possibly of partial double counting, but we conservatively list both effects separately. (b) Wavemeter offset: an offset of the He–Ne laser calibration source, estimated to be $\pm 3 \text{ MHz}$ at 730 nm by offline calibration. (c) 730-nm cavity resonance-frequency correction: the frequency of the generated 730-nm pulse was measured to be shifted from that of the continuous-wave 730-nm seed laser. This shift of about 10 MHz at 730 nm was regularly monitored, and was corrected for in our frequency determination. We conservatively assign an error of $10/\sqrt{12} = 3 \text{ MHz}$ to this correction at 730 nm (the standard deviation of a uniform distribution with a width of 10 MHz). (d) Residual $1S_c$ state contamination: our earlier studies with shorter running times^{11,22} indicate there is a residual population of the order of 5% of the $1S_c$ state after the microwave-driven clearing procedure, which was corrected for in the analysis above. We estimate the error in this correction by analysing the data assuming no residual $1S_c$ population. We take 68% of the differences between the two analysis results (33.5 MHz and 24 MHz for the $2P_c$ and $2P_f$ transitions, respectively) as 1σ uncertainties in the correction. (e) Magnetic field: the field at the magnetic minimum of the ALPHA-2 trap was measured in situ using the electron cyclotron resonance (ECR) method³⁵. A conservative uncertainty of 10 MHz in the ECR measurement gives a B field error of $3.6 \times 10^{-4} \text{ T}$, which in turn gives frequency errors of 5 MHz and 8 MHz for the 1S– $2P_c$ and 1S– $2P_f$ transitions, respectively, at 1 T. We take these values as a

measure of the uncertainty due both to the absolute value and to the run-to-run stability of the B field. We note that the frequency uncertainty in the 1S–2S transition due to B -field variations is negligible for our purposes¹¹. (f) Statistical uncertainties of the fit: these represent statistical uncertainties in the fit both from the experimental data and from the simulations. (g) Model uncertainties: described above.

The total errors for each transition are given by the quadratic sum of errors (a)–(g). Care must be taken when taking an average or a difference of the transition frequencies. Here we assume that error (b), the wavemeter offset, introduces a common offset to all the data series. The other errors are assumed to be uncorrelated across the dataset. The resulting combined uncertainty for the transition frequencies of antihydrogen is 39 MHz or 16 ppb (Fig. 4, average value). We expect that virtually all of the uncertainties can be considerably reduced in the near future owing to increased statistics and improved control of the systematics.

Determination of the fine-structure splitting and the Lamb shift of antihydrogen

To analyse the Zeeman-shifted energy levels of antihydrogen in the 2P state, we used the following Hamiltonian for the 2P state, which includes the field-free Hamiltonian (\hat{H}_0), the fine-structure Hamiltonian (\hat{H}_{fs}), the Zeeman Hamiltonian (\hat{H}_Z) and the hyperfine-structure Hamiltonian (\hat{H}_{hf}):

$$\hat{H} = \hat{H}_0 + \hat{H}_{fs} + \hat{H}_Z + \hat{H}_{hf} \quad (1)$$

$$\hat{H}_{fs} = \frac{2}{3} \bar{\mathcal{E}}_{fs} \left(\frac{1}{\hbar^2} \mathbf{L}_{\bar{e}} \cdot \mathbf{S}_{\bar{e}} + 1 \right) \quad (2)$$

$$\hat{H}_Z = -\frac{2\mu_{\bar{e}}(2P)}{\hbar} \mathbf{S}_{\bar{e}} \cdot \mathbf{B} - \frac{2\mu_{\bar{p}}}{\hbar} \mathbf{I}_{\bar{p}} \cdot \mathbf{B} + \frac{\tilde{\mu}_B}{\hbar} \mathbf{L}_{\bar{e}} \cdot \mathbf{B} \quad (3)$$

$$\hat{H}_{hf} = \frac{\bar{C}_{IL}}{\hbar^2} \mathbf{I}_{\bar{p}} \cdot \mathbf{L}_{\bar{e}} + \frac{\bar{C}_{IS}}{\hbar^2} [\mathbf{I}_{\bar{p}} \cdot \mathbf{S}_{\bar{e}} - 3(\mathbf{I}_{\bar{p}} \cdot \mathbf{r})(\mathbf{S}_{\bar{e}} \cdot \mathbf{r})] \quad (4)$$

Here, $\mathbf{L}_{\bar{e}}$ is the orbital angular momentum of the positron, $\mathbf{S}_{\bar{e}}$ is the spin angular momentum of the positron, $\mathbf{I}_{\bar{p}}$ is the nuclear spin angular momentum of the antiproton and \mathbf{r} is the position vector of the positron. $\bar{\mathcal{E}}_{fs}$ is the fine-structure splitting of antihydrogen at zero field. The magnetic moments of the positron and antiproton are given by $\mu_{\bar{e}}(2P) = \frac{|\bar{g}_s|}{2} \frac{|\bar{e}| \hbar}{2m_{\bar{e}}} \left(1 - \frac{\alpha^2}{10} \right)$ and $\mu_{\bar{p}} = -\frac{|\bar{g}_p|}{2} \frac{|\bar{e}| \hbar}{2m_{\bar{p}}}$ where \bar{g}_s and \bar{g}_p are the positron spin and antiproton g -factors, respectively, \bar{e} and $m_{\bar{e}}$ are the charge and mass of the positron, correspondingly, and α is the fine-structure constant. The last term of equation (4) is the Zeeman interaction due to the orbital angular momentum of the positron with magnetic moment of $\tilde{\mu}_B = -\left(1 - \frac{m_{\bar{e}}}{m_{\bar{p}}} \right) \frac{|\bar{e}| \hbar}{2m_{\bar{e}}}$, where $m_{\bar{p}}$ is the mass of the antiproton. \bar{C}_{IL} is the hyperfine-coupling constant due to the antiproton spin and the orbital angular momentum of the positron, and \bar{C}_{IS} is the hyperfine interaction due to the magnetic dipole–dipole interaction.

For the analysis of the classic Lamb shift ($\bar{\mathcal{E}}_{Lamb}$) and the fine-structure ($\bar{\mathcal{E}}_{fs}$) parameters of antihydrogen, we assumed that the absolute values of the three magnetic moments ($\mu_{\bar{e}}$, $\mu_{\bar{p}}$ and $\tilde{\mu}_B$) are the same as those of hydrogen. Previous measurements of the basic properties of antiparticles are consistent with this assumption. The hyperfine-coupling constants are also assumed to be those of hydrogen³⁸, $\bar{C}_{IL} = 22.2$ MHz and $\bar{C}_{IS} = -22.2$ MHz.

Our measurements determine the energy levels, with respect to the 1S ground state, of two of the Zeeman sublevels in the $n = 2$ positronic manifold of antihydrogen at a magnetic field of 1.0329 T. Specifically, the $2P_f$ state belongs to the $2P_{1/2}$ manifold, and the $2P_c$ state belongs to the $2P_{3/2}$ manifold (see Fig. 1). We combine these results with our previous measurement of the $1S_d$ – $2S_d$ transition⁷ and assume the validity of

the standard Zeeman and hyperfine interactions to derive the fine-structure splitting $\bar{\mathcal{E}}_{fs}$ (that is, the energy difference between $2P_{1/2}$ and $2P_{3/2}$), and the classic Lamb shift $\bar{\mathcal{E}}_{Lamb}$ (that is, the energy difference between $2S_{1/2}$ and $2P_{1/2}$), both defined at zero field.

Taking into account the hyperfine splitting, we find the energy separation between the $2P_c$ - and $2P_f$ -levels at 1.0329 T to be 14.945 ± 0.0975 GHz, from the difference of the weighted average values of the observed transition frequencies. Furthermore, we obtain the separation between the $2S_d$ and $2P_c$ -levels to be $\Delta E(2S, 2P_c) = 9,832 \pm 49$ MHz, and that between the $2S_d$ and $2P_f$ -levels to be $\Delta E(2S, 2P_f) = 24,778 \pm 60$ MHz, in the same field. The sum and the difference of the two quantities, $\Delta E(2S, 2P_c)$ and $\Delta E(2S, 2P_f)$, can be expressed by the following equations, which are based on the standard Hamiltonian of the hydrogen atom in a magnetic field B (refs.^{36,37}). We neglect terms that contribute less than 1 MHz.

$$\Delta E(2S, 2P_f) - \Delta E(2S, 2P_c) = 2E_1(-B) + \frac{1}{2} \bar{C}_{IL} \cos(2\sigma) + \frac{1}{10} [\cos(2\sigma) - 3\sqrt{2} \sin(2\sigma)] \bar{C}_{IS} \quad (5)$$

$$\Delta E(2S, 2P_f) + \Delta E(2S, 2P_c) = 2\bar{\mathcal{E}}_{Lamb} - 2\bar{\mathcal{E}}_{fs} + \frac{1}{2} \bar{C}_{IL} - \frac{3}{10} \bar{C}_{IS} + \frac{1}{2} \bar{\mathcal{E}}_{hf}(2S) - [2\mu_{\bar{e}}(2P) + \tilde{\mu}_B] B \quad (6)$$

where

$$E_1(-B) = \sqrt{\left\{ \frac{1}{6} \bar{\mathcal{E}}_{fs} - B \left[\mu_{\bar{e}}(2P) + \frac{1}{2} \tilde{\mu}_B \right] \right\}^2 + \frac{2}{9} \bar{\mathcal{E}}_{fs}^2}$$

and

$$\tan \sigma = \frac{-\bar{\mathcal{E}}_{fs} + [6\mu_{\bar{e}}(2P) + 3\tilde{\mu}_B] B + 6E_1(-B)}{2\sqrt{2} \bar{\mathcal{E}}_{fs}}$$

Here, $\bar{\mathcal{E}}_{hf}(2S)$ is the hyperfine splitting in the 2S state at zero field.

Finally, using the CODATA 2014 values of the fundamental constants for the hydrogen atom³⁹, the fine-structure splitting $\bar{\mathcal{E}}_{hf}$ and the classic Lamb shift $\bar{\mathcal{E}}_{Lamb}$ of the antihydrogen atom are determined by numerically solving equations (5) and (6) with the measured energy-level differences given in Table 2 as input.

Hydrogen transition frequencies in a magnetic field

From zero-field measurements in hydrogen for the $1S_{1/2}$ – $2S_{1/2}$ (ref.⁴⁰), $2S_{1/2}$ – $2P_{1/2}$ (ref.⁴¹) and $2P_{1/2}$ – $2P_{3/2}$ (ref.⁴²) transitions, we obtain hyperfine centroid frequencies of

$$1S-2P_{1/2} \text{ transition: } 2,466,060,355 \text{ MHz}$$

$$1S-2P_{3/2} \text{ transition: } 2,466,071,324 \text{ MHz}$$

The transition frequencies at 1.0329 T (Table 2) are calculated by evaluating corrections assuming the standard Zeeman, fine-structure and hyperfine interactions in a magnetic field^{36,37} and using the current CODATA values of the fundamental constants³⁹. The precision of our calculations is better than 1 MHz.

In comparing the hydrogen values with the measured antihydrogen frequencies in Table 2 and Fig. 4, the value of the magnetic field was assumed to be exact for the hydrogen case.

Data availability

The datasets generated and/or analysed during the current study are available from J.S.H. on reasonable request.

35. Amole, C. et al. In situ electromagnetic field diagnostics with an electron plasma in a Penning–Malmberg trap. *New J. Phys.* **16**, 013037 (2014).

36. Rasmussen, C. Ø., Madsen, N. & Robicheaux, F. Aspects of 1S–2S spectroscopy of trapped antihydrogen atoms. *J. Phys. B* **50**, 184002 (2017); corrigendum **51** 099501 (2018).

37. Melezhik, V. S. & Schmelcher, P. Quantum energy flow in atomic ions moving in magnetic fields. *Phys. Rev. Lett.* **84**, 1870–1873 (2000).
38. Kramida, A. E. A critical compilation of experimental data on spectral lines and energy levels of hydrogen, deuterium, and tritium. *At. Data Nucl. Data Tables* **96**, 586–644 (2010); erratum **126**, 295–298 (2019).
39. Mohr, P. J., Newell, D. B. & Taylor, B. N. CODATA recommended values of the fundamental physical constants: 2014. *Rev. Mod. Phys.* **88**, 035009 (2016).
40. Parthey, C. G. et al. Improved measurement of the hydrogen 1S–2S transition frequency. *Phys. Rev. Lett.* **107**, 203001 (2011).
41. Lundeen, S. R. & Pipkin, F. M. Measurement of the Lamb shift in hydrogen $n = 2$. *Phys. Rev. Lett.* **46**, 232–235 (1981).
42. Hagley, E. W. & Pipkin, F. M. Separated oscillatory field measurement of hydrogen $2S_{1/2}$ – $2P_{3/2}$ fine structure interval. *Phys. Rev. Lett.* **72**, 1172–1175 (1994).

Acknowledgements This work was supported by: the European Research Council through its Advanced Grant programme (J.S.H.); CNPq, FAPERJ, RENAFAP (Brazil); NSERC, ALPHA-g/CRUCS CFI, NRC/TRIUMF, EHPDS/EHDRS (Canada); FNU (Nice Centre), Carlsberg Foundation (Denmark); ISF (Israel); STFC, EPSRC, the Royal Society and the Leverhulme Trust (UK); DOE, NSF (USA); and VR (Sweden). We are grateful for the efforts of the CERN AD team, without which these experiments could not have taken place. We thank P. Djuricanin (University of British Columbia) for extensive help with the laser system and calibrations. We thank J. Tonoli and the CERN staff, as well as T. Mittertreiner and the UBC staff, for extensive, time-critical help with machining and electrical works. We thank the staff of the Superconducting Magnet Division at Brookhaven National Laboratory for collaboration and for the fabrication of the trapping magnets. We thank C. Marshall (TRIUMF) for work on the ALPHA-2 cryostat. We thank F. Besenbacher (Aarhus) for timely support in procuring the

ALPHA-2 external solenoid. We thank T. Miller (Ohio) for advice on the initial development of the pulsed laser system. We thank A. Kostelecky and G. Drake for discussions on the theoretical aspects of this work.

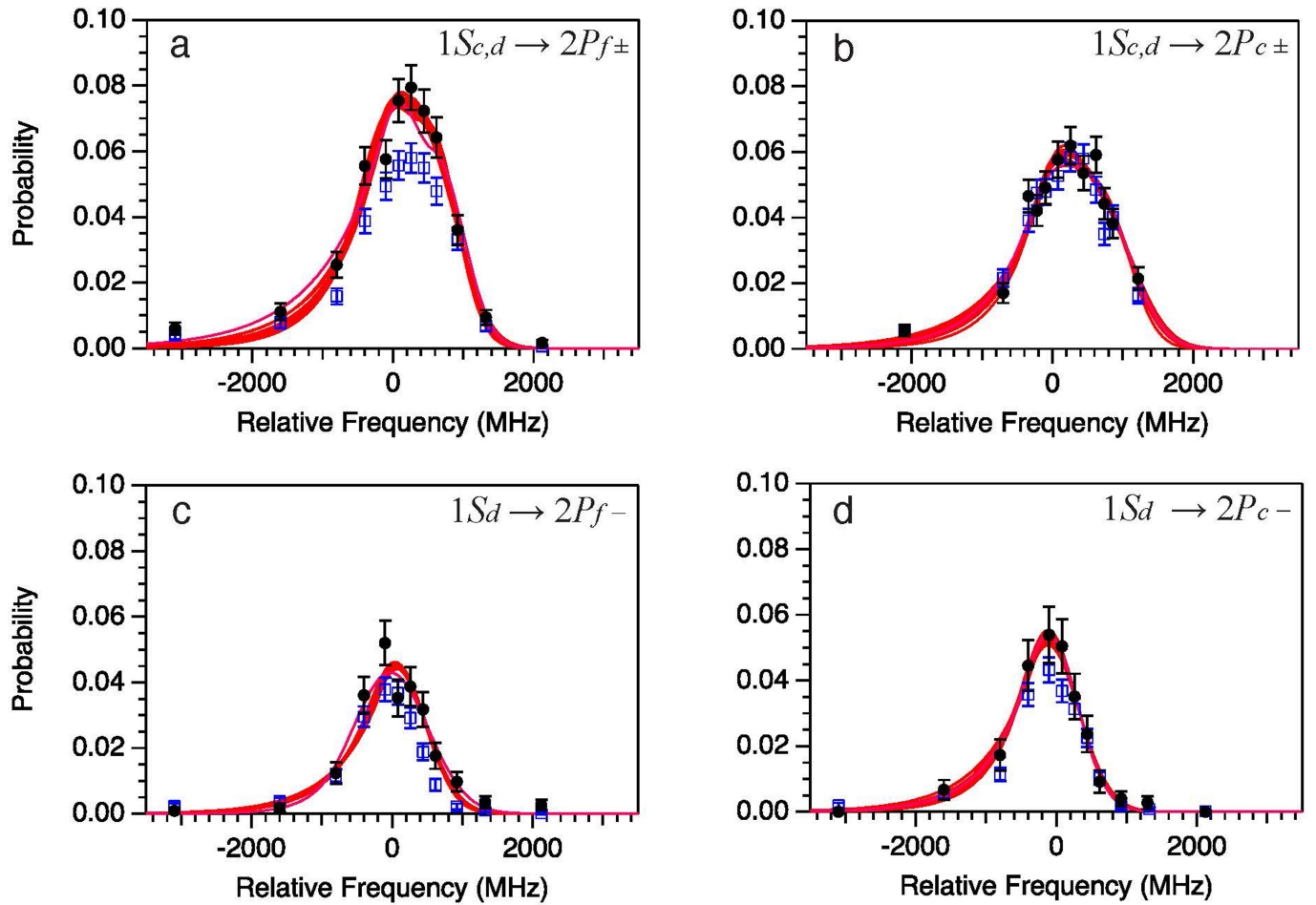
Author contributions This study was based on data collected using the ALPHA-2 antihydrogen trapping apparatus, designed and constructed by the ALPHA Collaboration using methods developed by the entire collaboration. The entire collaboration participated in the operation of the apparatus and the data-taking activities. Pulsed Lyman- α spectroscopy was first proposed by M.C.F. and developed further by T.M., F.R., J.M.M., R.C., A.E., A.K. and M.C.F. The original laser was designed by T.M. and tested by J.M.M. and T.M. The laser system at CERN was implemented, commissioned and operated by J.M.M., R.C., A.E., A.K. and T.M. The simulation program for laser interaction with magnetically trapped atoms was developed by F.R. The microwave techniques and ECR magnetometry were developed by T.F., M.E.H. and W.N.H. The positron accumulator is the responsibility of C.J.B., M.C., M.S., C.A.I. and D.P.v.d.W. The annihilation detector system was developed by A.C., M.C.F., D.R.G., L.K., J.T.K.M., S.M., K.O., A.O. and P.P. Detailed analysis of the antiproton annihilation detector data was done by J.T.K.M., A.E. and A.O. The frequency determination and the evaluation of its uncertainty was performed by T.M., M.C.F., A.K., A.E., R.C., R.I.T. and A.O. The manuscript was written by T.M., M.C.F. and J.S.H., with significant input from A.K., R.C., A.E., A.O., J.T.K.M., M.E.H., N.M., C.Ø.R. and F.R. The manuscript was edited and improved by the entire collaboration.

Competing interests The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.C.F., J.S.H. or T.M.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | Determination of transition frequencies. a–d, For each series, the experimental data (filled black circles with error bars) are plotted with fits of various models (red lines) discussed in Methods. The experimental data are normalized to the total number of the detected antihydrogen atoms and a laser power of 5 nW. Also shown are the results of standard simulations (open blue squares with error bars), similarly normalized to the total number of

simulated atoms, illustrating the degree of agreement between the data and the simulations, without any tuning parameters. Some discrepancies in the amplitudes can be observed, which may point to errors in our laser power estimates. We note that because our frequency-fitting procedure allows variations in the relative amplitudes, the fits are largely insensitive to the amplitude differences (Methods). Error bars represent 1σ .

Measurement of the quantum geometric tensor and of the anomalous Hall drift

<https://doi.org/10.1038/s41586-020-1989-2>

Received: 10 January 2019

Accepted: 10 December 2019

Published online: 19 February 2020

 Check for updates

A. Gianfrate^{1,4}, O. Bleu^{2,4}, L. Dominici¹, V. Ardizzone¹, M. De Giorgi¹, D. Ballarini¹, G. Lerario¹, K. W. West³, L. N. Pfeiffer³, D. D. Solnyshkov², D. Sanvitto^{1✉} & G. Malpuech^{2✉}

Topological physics relies on the structure of the eigenstates of the Hamiltonians. The geometry of the eigenstates is encoded in the quantum geometric tensor¹—comprising the Berry curvature² (crucial for topological matter)³ and the quantum metric⁴, which defines the distance between the eigenstates. Knowledge of the quantum metric is essential for understanding many phenomena, such as superfluidity in flat bands⁵, orbital magnetic susceptibility^{6,7}, the exciton Lamb shift⁸ and the non-adiabatic anomalous Hall effect^{6,9}. However, the quantum geometry of energy bands has not been measured. Here we report the direct measurement of both the Berry curvature and the quantum metric in a two-dimensional continuous medium—a high-finesse planar microcavity¹⁰—together with the related anomalous Hall drift. The microcavity hosts strongly coupled exciton–photon modes (exciton polaritons) that are subject to photonic spin–orbit coupling¹¹ from which Dirac cones emerge¹², and to exciton Zeeman splitting, breaking time-reversal symmetry. The monopolar and half-skyrmion pseudospin textures are measured using polarization-resolved photoluminescence. The associated quantum geometry of the bands is extracted, enabling prediction of the anomalous Hall drift, which we measure independently using high-resolution spatially resolved epifluorescence. Our results unveil the intrinsic chirality of photonic modes, the cornerstone of topological photonics^{13–15}. These results also experimentally validate the semiclassical description of wavepacket motion in geometrically non-trivial bands^{9,16}. The use of exciton polaritons (interacting photons) opens up possibilities for future studies of quantum fluid physics in topological systems.

One of the key manifestations of topological effects in physics is the conductance quantization in the two-dimensional (2D) quantum Hall effect. This perfect quantization relies on a topological invariant characterizing the global band properties: the Chern number. Non-zero Chern numbers are associated with the chiral conducting edge states in topological insulators and superconductors³. Beyond electronic systems, topological band concepts have been extended to various wave systems covering photonics^{14,15}, acoustics¹⁷, cold atoms¹⁸ and even geophysics¹⁹.

Topological effects are not encoded in the energy spectrum of the system but rely on the non-trivial geometry of the eigenstates. It is the gauge-invariant quantum geometric tensor (QGT) that contains the structural information about the eigenstates of a parametrized Hamiltonian. The QGT has a symmetric real part which defines the quantum metric characterizing distances between states⁴ in a parameter space. Its antisymmetric imaginary part determines the Berry curvature² whose momentum space distribution is crucial in modern Physics. Locally, the Berry curvature is responsible for the anomalous Hall transport¹⁶ in the intrinsic spin Hall and valley Hall effects.

Its integral over a 2D closed manifold gives the Chern number. Additionally, the quantum metric, related to the concept of fidelity in quantum information theory, also describes important physical phenomena. It can probe quantum phase transitions when defined in an arbitrary parameter space²⁰. The momentum-space metric affects the electronic orbital magnetic susceptibility^{6,7} in crystals and the exciton Lamb shift in transition-metal dichalcogenides⁸. It characterizes superfluidity and the current of bogolons (superfluid excitations) in flat bands^{5,21} and also corrections to the semiclassical equations describing the anomalous Hall effect^{6,9}.

The extension of topological concepts from solid state physics to other classical or quantum systems has opened up possibilities for measuring the local geometrical properties of bands, not just the global properties (such as the conductivity measured in the quantum Hall effect). Several protocols have been proposed to measure the Berry curvature^{22,23}. Experimental reconstructions via indirect dynamical measurements have been reported^{24,25}. The parameter space geometry of two-level systems has been explored experimentally even more recently^{26,27}. In this work, we present a direct measurement of the

¹CNR NANOTEC, Istituto di Nanotecnologia, Lecce, Italy. ²Institut Pascal, PHOTON-N2, Université Clermont Auvergne, CNRS, SIGMA Clermont, Clermont-Ferrand, France. ³PRISM (Princeton Institute for the Science and Technology of Materials), Princeton University, Princeton, NJ, USA. ⁴These authors contributed equally: A. Gianfrate, O. Bleu. ✉e-mail: daniele.sanvitto@cnr.it; guillaume.malpuech@uca.fr

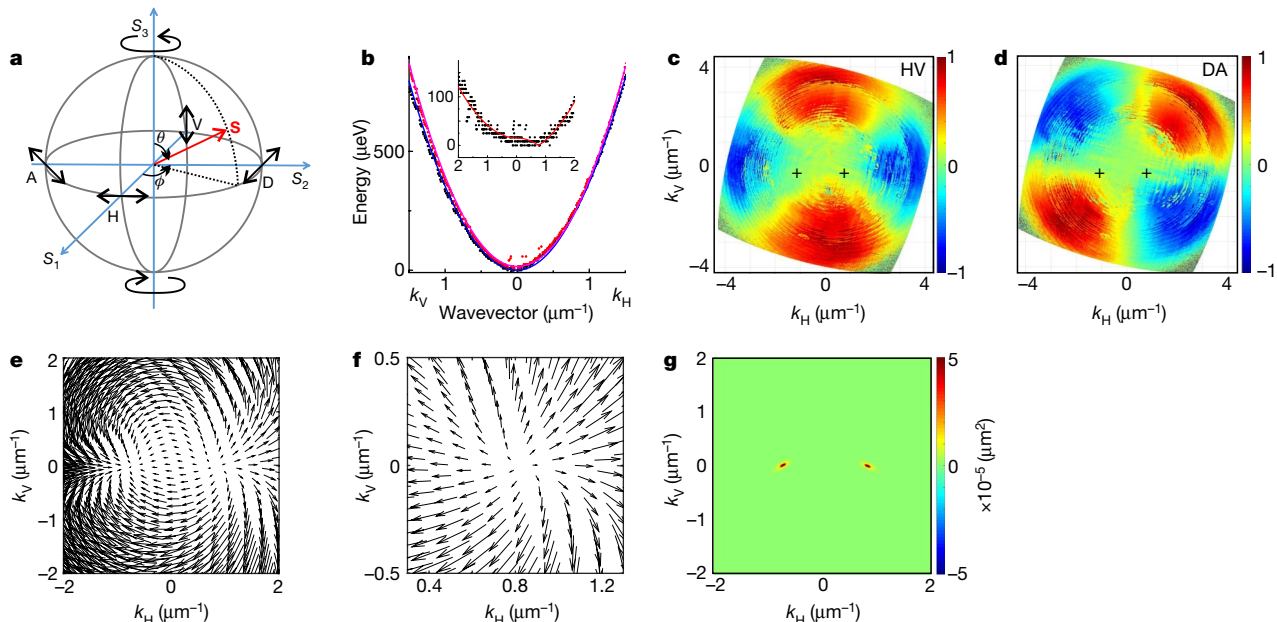


Fig. 1 | Emergence of pseudospin monopoles from photoluminescence at OT. **a**, Pseudospin (\mathbf{S}) orientation on the Poincaré sphere, parametrized by θ and ϕ and given by the emission polarization degree on the HV, DA and circular RL basis; equation (5). **b**, Eigenmode energies (zero at about 1,601.5 meV) along k_H and k_V . Inset, energy splitting of eigenmodes. Points, experiment; pink and blue lines, dispersions of the two branches fitted with equation (1). **c**, **d**, Degree

of polarization of the lower mode in k space for HV (**c**) and DA (**d**). The crosses mark the degeneracy points. **e**, k -space in-plane pseudospin (S_1, S_2) texture. **f**, A magnification of one of the crossing points that demonstrates a monopole texture. The lengths of the arrows in **e**, **f**, are in arbitrary units. **g**, Quantum metric tensor trace ($g_{HH} + g_{VV}$), with peak around the monopoles.

full momentum space QGT (Berry curvature and quantum metric) of the 2D bands of a homogeneous system. Furthermore, we measure independently the dynamics of an accelerated wavepacket which demonstrates the anomalous Hall drift. This drift is reproduced by the semiclassical equations of motion^{9,16} including the measured band geometry as an input parameter.

Our experimental platform is a high-quality planar microcavity (quality factor, $Q > 10^5$) with embedded quantum wells that support 2D strongly coupled exciton–photon bands (see Supplementary Fig. 6)¹⁰. Each band is doubly polarization degenerate and forms a pseudo-spinor. The polarization degeneracy is lifted by the photonic splitting between the transverse electric (TE) and transverse magnetic (TM) eigenmodes¹¹ and, under magnetic field, by the exciton Zeeman splitting. The polarized polariton eigenstates are exactly determined by a Fourier space mapping of polarization-resolved photoluminescence. They exhibit non-zero Berry curvature and a non-zero quantum metric as discussed below.

Theoretical background

Before presenting the measurements, let us recall the properties of an effective two-band Hamiltonian²⁸ describing, in the conservative limit, the lower polariton branch in the circular polarization basis¹⁰:

$$H_{\mathbf{k}} = \begin{pmatrix} \frac{\hbar^2 k^2}{2m^*} + \Delta_z & \alpha - \beta k^2 e^{-2i\varphi} \\ \alpha - \beta k^2 e^{2i\varphi} & \frac{\hbar^2 k^2}{2m^*} - \Delta_z \end{pmatrix} \quad (1)$$

where $m^* = m_l m_t / (m_l + m_t)$ and m_l and m_t are the longitudinal and transverse effective masses; $k = |\mathbf{k}| = \sqrt{k_H^2 + k_V^2}$ is the in-plane wavevector ($k_H = k \cos \varphi$, $k_V = k \sin \varphi$, and φ is the propagation angle); Δ_z is the polariton Zeeman splitting owing to the excitonic part; and α is the optical birefringence, unavoidable in crystalline systems. The birefringence

leads to a k -independent splitting between horizontally and vertically polarized states. β quantifies the \mathbf{k} -dependent TE–TM splitting, ubiquitous in 2D photonic systems, which makes 2D photonic bands geometrically non-trivial; and \hbar is the reduced Planck's constant. This 2×2 Hamiltonian can be decomposed into a linear combination of Pauli matrices, describing the interaction between an effective magnetic field and a pseudospin:

$$H_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m^*} \mathbf{I} + \mathbf{\Omega}(\mathbf{k}) \cdot \boldsymbol{\sigma} \quad (2)$$

where $\boldsymbol{\sigma}$ is a vector of Pauli matrices (the spin operator), the average value of which is $\mathbf{S} = \langle \boldsymbol{\sigma} \rangle$. \mathbf{S} is here the polarization pseudospin of light. The effective field $\mathbf{\Omega}(\mathbf{k})$ reads:

$$\mathbf{\Omega}(\mathbf{k}) = \begin{pmatrix} \alpha - \beta k^2 \cos 2\varphi \\ -\beta k^2 \sin 2\varphi \\ \Delta_z \end{pmatrix} \quad (3)$$

The pseudospins of the eigenstates are parallel and antiparallel with the effective field. The QGT components g_{ij} and B_z are linked to the variation of the pseudospin orientation in k space as²⁸:

$$g_{ij} = \frac{1}{4} \left(\partial_{k_i} \theta \partial_{k_j} \theta + \sin^2 \theta \partial_{k_i} \phi \partial_{k_j} \phi \right) \quad (4a)$$

$$B_z = \frac{1}{2} \sin \theta \left(\partial_{k_i} \theta \partial_{k_j} \phi - \partial_{k_j} \theta \partial_{k_i} \phi \right) \quad (4b)$$

with g_{ij} the metric components and B_z the Berry curvature in the z direction, which cancel if the TE–TM spin-orbit coupling vanishes ($\beta = 0$). $\theta(\mathbf{k})$ and $\phi(\mathbf{k})$ are, respectively, the polar and azimuthal angles parametrising the eigenstate $\psi = [\cos(\theta/2)e^{-i\varphi}, \sin(\theta/2)]^T$ (where T indicates the transpose) and the pseudospin position on

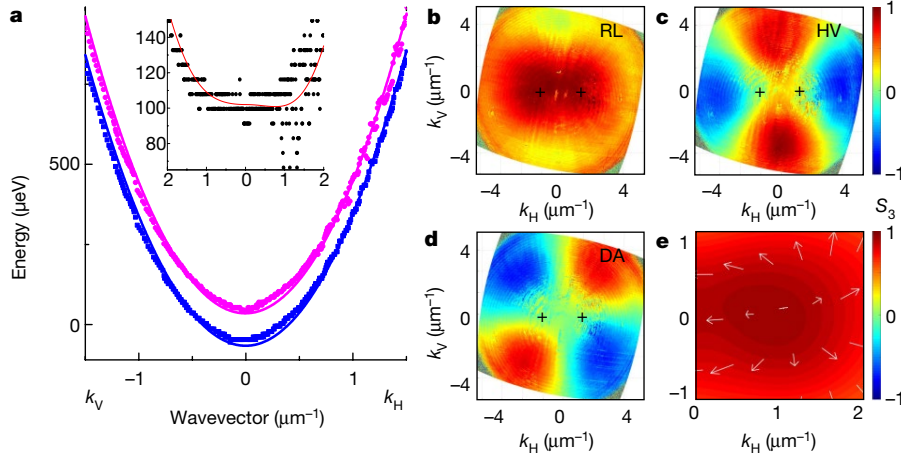


Fig. 2 | Broken time-reversal symmetry: emergence of half-skyrmion pseudospin textures, from photoluminescence at 9 T. **a**, Energy dispersion along k_H and k_V (zero at about 1,601.9 meV). Anticrossing of the branches is observed instead of their crossing. The polarization bands are split for all wavevectors (see inset, where along k_H the splitting has a non-zero minimum). The pink and blue lines are the fits of the energy dispersions. **b–d**, Polarization-

degree maps of the lower energy mode for RL (**b**), HV (**c**) and DA (**d**). The crosses mark the anticrossing points. **e**, Pseudospin distribution in k -space, magnified near one of the crossing points. The in-plane pseudospin (S_1, S_2) is shown by the white arrows (lengths in arbitrary units), and the S_3 amplitude (dimensionless) is shown by colour.

the Poincaré sphere (Fig. 1a) with $\theta = \arccos S_3$ and $\phi = \arctan (S_2/S_1)$, where S_1, S_2, S_3 are the components of the pseudospin vector \mathbf{S} . The QGT components g_{ij} and B_z are computed analytically²⁸ in Supplementary Note 1 using the eigenstates of equation (1) with $\mathbf{S} \parallel \mathbf{\Omega}$.

Quantum geometry of emergent Dirac cones

The sample studied is a high-quality microcavity with a 100-ps lifetime (see Methods section ‘Sample details’). The experimental setup is shown in Extended Data Fig. 1. The measurements are executed at 4 K in a reflection configuration under an applied external magnetic field. We first use off-resonant continuous-wave laser excitation. The photoluminescence is measured versus the 2D wavevector and energy for all six polarization axes of the Poincaré sphere (Fig. 1a) corresponding to the horizontal–vertical (HV), diagonal–antidiagonal (DA), and circular right–left (RL) polarizations. For each wavevector, the energies of the polarization doublet are found by Gaussian fitting of the emission. Their pseudospin (Stokes vector) is determined from the polarization intensities as:

$$S_1(\mathbf{k}) = \frac{I_H - I_V}{I_H + I_V}, \quad S_2(\mathbf{k}) = \frac{I_D - I_A}{I_D + I_A}, \quad S_3(\mathbf{k}) = \frac{I_R - I_L}{I_R + I_L} \quad (5)$$

The k -space pseudospin distribution then enables us to compute the QGT components of each branch using equation (4).

Figure 1 shows the 0 T measurements (no Zeeman splitting, $\Delta_z = 0$) at zero exciton–photon detuning (see Methods section ‘Exciton–photon detuning’). The energy dispersion extracted from the raw photoluminescence (see Methods section ‘Experimental setup to measure the QGT’ and Supplementary Note 2) is shown in Fig. 1b. The inset shows the energy difference between the modes. By fitting the dispersion, we obtain the polariton mass $m = (9.2 \pm 0.1) \times 10^{-5} m_0$ (m_0 is the free electron mass), the TE–TM splitting $2\beta = 26.3 \pm 0.3 \mu\text{eV} \mu\text{m}^2$, and the birefringence (HV splitting) $2\alpha = 15.3 \pm 0.6 \mu\text{eV}$ at $k = 0$. (Errors denote the standard deviation of the fitting parameters.) If the HV splitting were zero ($\alpha = 0$), the dispersion would be composed of two parabola of different masses touching at $k = 0$, similar to the quadratic band degeneracies in bilayer graphene. The Berry phase accumulated along a closed loop around the band touching point would be 2π (a Berry topological charge of 1). When the HV splitting is non-zero ($\alpha \neq 0$), as

in our sample, the cylindrical symmetry is broken. Along k_H , the lowest energy mode has the smallest mass. The two parabola cross at $k_0 = \sqrt{\alpha/\beta} \approx 0.8 \mu\text{m}^{-1}$, where the TE–TM and HV splitting cancel each other. Along k_V such points are absent, because both contributions have the same sign. This is visible in Fig. 1c, d, which shows the HV and DA polarization degree of the lower band (the circular polarization degree is zero at 0 T).

The degeneracy points, marked by crosses, are tilted Dirac cones, around which the effective field and pseudospin textures appear similar to 2D monopoles (Fig. 1e, f). The breaking of the TE–TM rotational symmetry by the HV field induces the separation of the TE–TM vector field of winding number 2 into a pair of 2D monopoles of winding 1, but of opposite divergences. Each monopole carries a Berry topological charge of 1/2, so that the band topology does not depend on the HV splitting, but the band geometry does. The Berry curvature of each monopole is a delta function, whereas the metric has a finite extension measured in Fig. 1g. Therefore, any finite-duration measurement of the Berry phase performed by making a loop around these points should show a deviation from the adiabatic value of π quantified by this metric distribution. These effective monopoles can be mapped to emergent non-Abelian gauge fields acting on photons¹². Interestingly, the metric distribution around the crossing points is not cylindrically symmetric, which might be due to non-hermiticity²⁹.

Now we break the time-reversal symmetry (Fig. 2), applying a 9 T magnetic field described by the Zeeman term Δ_z . The field also makes the exciton–photon detuning slightly negative (see Methods section ‘Exciton–photon detuning’), owing to the exciton diamagnetic shift. Figure 2a shows the dispersions along k_H and k_V , as in Fig. 1b (the inset shows the energy difference). The modes are now split everywhere in k space, and at $k = 0$ the energy splitting is approximately 102 μeV . The crossing along k_H becomes an anticrossing. The splitting at the anticrossing point is the polariton Zeeman splitting $2\Delta_z = 100.9 \pm 0.6 \mu\text{eV}$ (where the error is the standard deviation of the parameter), caused by the excitonic part (exciton g factor of about 0.2). It is much larger (by ten times) than the linewidth of our ultrahigh-quality sample, despite the optical frequency operation. The measured polarization degrees are shown in Fig. 2b–d. The polarization at $k = 0$ becomes elliptical. The circular polarization degree decreases along k_V and increases along k_H up to $\pm k_0$, where it becomes close to 1. A magnification of the measured pseudospin texture around k_0 is shown in Fig. 2e, exhibiting a part of a half-skyrmion, as expected.

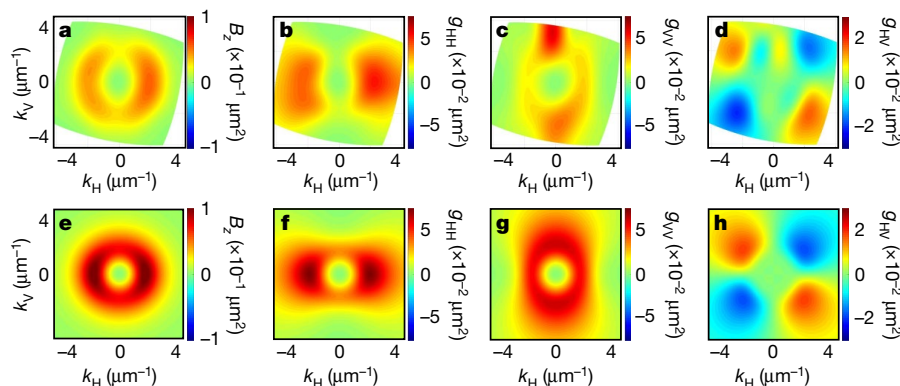


Fig. 3 | Berry curvature and quantum metric distributions. **a–d**, Experiment, k -space distribution of QGT elements (from photoluminescence at 9 T): Berry curvature B_z (**a**); g_{HH} (**b**); g_{VV} (**c**); and g_{HV} (**d**); extracted using equation (4). **e–h**,

Theory. The computations are based on the effective Hamiltonian (equation (1)) and Supplementary Note 1.

The k -space distributions of the Berry curvature and of the three components of the quantum metric tensor, extracted from the experimental data of Fig. 2 using equation (4), are shown in Fig. 3a–d. They are compared with analytical calculations (Fig. 2e–h; formula given in Supplementary Note 1) performed using the parameters extracted from the dispersions in Figs. 1b, 2a. Without HV splitting, the Berry curvature would be circularly symmetric, whereas for a dominating HV splitting the distribution would be concentrated around the anticrossing points. Here, we are between these two limiting cases. A similar procedure applied to the second polarization branch (see Supplementary Fig. 4), confirms that the two branches are cross-polarized, show opposite Berry curvatures and the same quantum metric elements.

Anomalous Hall effect

A consequence of non-trivial band geometry is the anomalous Hall drift of an accelerated wavepacket which appears as a correction in the semiclassical equation of motion¹⁶:

$$\hbar \frac{\partial \mathbf{r}}{\partial t} = \frac{\partial E}{\partial \mathbf{k}} + \mathbf{F} \times \mathbf{B} \quad (6)$$

where \mathbf{r} is the centre of mass of the wavepacket, $E(\mathbf{k})$ is the dispersion, $\mathbf{F}(\mathbf{k})$ is the accelerating force and $\mathbf{B} = B_z \mathbf{e}_z$, where \mathbf{e}_z is the unit vector in the z direction. The acceleration is provided by the thickness gradient of the microcavity. The resulting energy gradient accelerates polaritons similar to the way an electric field accelerates charges. We choose a sample region with the largest gradients of 6 meV mm^{−1} and negative exciton–photon detuning. The gradient is measured independently (see Supplementary Note 3) and exhibits a slight spatial variation (a saddle-type potential). We selectively excite the lower polariton branch at $k = 0$ with a continuous-wave laser (30-μm-diameter spot). Figure 4a shows the spatial distribution of the intensity at ± 9 T (removing noise) under elliptically polarized excitation of the lower eigenstate (see Methods section ‘Experimental setup to measure the polariton anomalous Hall effect’). The two traces separate along their propagation. This is confirmed in Fig. 4b, which shows the measured centre-of-mass trajectories, well reproduced by numerical simulations based on the semiclassical equation (6), using as input parameters the potential and the Berry curvature distribution computed using equations (1) and (4) and the experimentally fitted parameters. Interestingly, this saddle-like potential magnifies the drift by a factor of 1.6 with respect to a constant gradient (see Supplementary Note 3). The oscillations of experimental trajectories are attributed to sample disorder. They remain smaller than the global drift. The role of non-adiabaticity on the trajectories,

quantified by the quantum metric⁹, cannot be evidenced here, owing to the experimental uncertainties. However, non-adiabaticity can be increased by modifying the excitation conditions. Figure 4c shows the cross-polarized emission when the excitation is circularly polarized, slightly exciting the upper polarization eigenstate (black, experiment; blue, full Schrödinger simulation presented in Supplementary Note 3). The contrast of these oscillations is the non-adiabatic fraction, also given by the distance between the quantum states at $k = 0$ and $k = k_0$ determined from the metric g_{ij} (more details given in Supplementary Note 4) as:

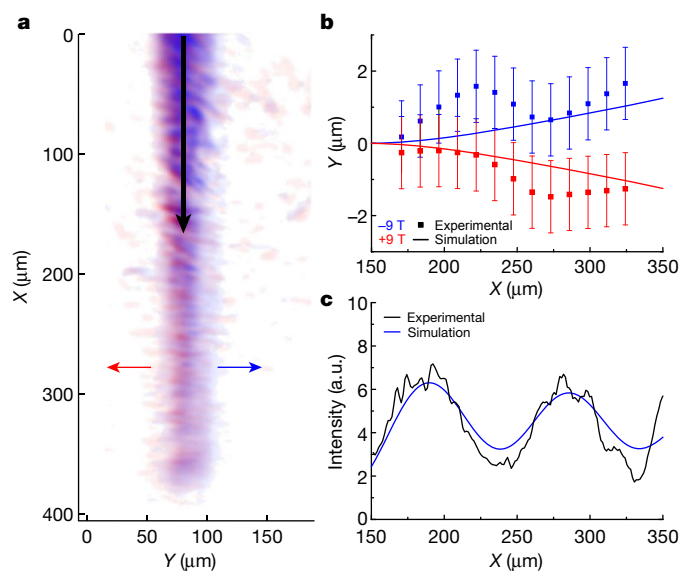


Fig. 4 | Polariton anomalous Hall effect. Continuous-wave-laser resonant excitation of the lower polariton mode at $k = 0$ with 30-μm-diameter excitation spot. **a**, Spatial distribution of emission at ± 9 T (pink and blue). **b**, Centre-of-mass trajectories at ± 9 T (red and blue). The experimental data (averaged over groups of points) are shown by squares with error bars representing the average value of the standard deviation within each group and the theory-based data from the semiclassical equation are shown by solid lines. In **a** and **b** the polarization is elliptical, corresponding to the polarization of the eigenstate. **c**, The polarization of the excitation is circular, which leads to oscillations in the intensity of the emission in the cross-circular polarization. Black, experiment; blue, theory. a.u., arbitrary units. X and Y are the in-plane axes of the sample.

$$|f_1|^2 \approx \left(\int_{|\psi_0(0)\rangle}^{|\psi_0(k_0)\rangle} \sqrt{\sum_{ij} g_{ij} \frac{dk^i}{dt} \frac{dk^j}{dt}} dt \right)^2 \quad (7)$$

where dt is taken along the shortest distance path. The distance that we estimate from the oscillations is 0.16 ± 0.02 , whereas equation (7) gives 0.18 ± 0.03 , showing a remarkable agreement. (The errors here correspond to the uncertainty estimates discussed in the Supplementary Information.)

Our experiments provide a measure of both the full non-trivial band geometry of a 2D continuous system and, independently, real-space wavepacket motion demonstrating anomalous Hall effect and non-adiabaticity. The experiments support the validity of the semiclassical approach and band geometry to compute wavepacket dynamics. Our results demonstrate that 2D photonic modes, because they are TE and TM polarized, carry topological charges, which is essential for topological photonics. Indeed, by using an appropriate lattice, the geometrically non-trivial bulk photon dispersion transforms into gapped topologically non-trivial photonic Bloch bands^{13–15,30,31} the QGT of which can be explored by our technique²⁸. The polaritonic platform (interacting photons) able to demonstrate lasing and quantum fluid behaviour (superfluidity, quantized vortices, etc.) opens up opportunities for topological physics. The platform has already enabled notable advances, such as topological lasers³², and offers exciting possibilities, such as mixing different topological effects related to quantum vortices and band structures³³.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-1989-2>.

- Berry, M. The quantum phase, five years after. In *Geometric Phases in Physics* (eds Wilczek, F. & Shapere, A.) 7–28 (World Scientific, 1989).
- Berry, M. V. Quantal phase factors accompanying adiabatic changes. *Proc. R. Soc. Lond. A* **392**, 45–57 (1984).
- Hasan, M. Z. & Kane, C. L. Topological insulators. *Rev. Mod. Phys.* **82**, 3045–3067 (2010).
- Provost, J. & Vallee, G. Riemannian structure on manifolds of quantum states. *Commun. Math. Phys.* **76**, 289–301 (1980).
- Peotta, S. & Törmä, P. Superfluidity in topologically nontrivial flat bands. *Nat. Commun.* **6**, 8944 (2015).
- Gao, Y., Yang, S. A. & Niu, Q. Field induced positional shift of Bloch electrons and its dynamical implications. *Phys. Rev. Lett.* **112**, 166601 (2014).

- Piéchon, F., Raoux, A., Fuchs, J.-N. & Montambaux, G. Geometric orbital susceptibility: quantum metric without Berry curvature. *Phys. Rev. B* **94**, 134423 (2016).
- Srivastava, A. & Imamoglu, A. Signatures of Bloch-band geometry on excitons: nonhydrogenic spectra in transition-metal dichalcogenides. *Phys. Rev. Lett.* **115**, 166802 (2015).
- Bleu, O., Malpuech, G., Gao, Y. & Solnyshkov, D. D. Effective theory of nonadiabatic quantum evolution based on the quantum geometric tensor. *Phys. Rev. Lett.* **121**, 020401 (2018).
- Kavokin, A., Baumberg, J. J., Malpuech, G. & Laussy, F. P. (eds) *Microcavities* (Oxford Univ. Press, 2011).
- Kavokin, A., Malpuech, G. & Glazov, M. Optical spin Hall effect. *Phys. Rev. Lett.* **95**, 136601 (2005).
- Terças, H., Flayac, H., Solnyshkov, D. D. & Malpuech, G. Non-Abelian gauge fields in photonic cavities and photonic superfluids. *Phys. Rev. Lett.* **112**, 066402 (2014).
- Haldane, F. D. M. & Raghu, S. Possible realization of directional optical waveguides in photonic crystals with broken time-reversal symmetry. *Phys. Rev. Lett.* **100**, 013904 (2008).
- Lu, L., Joannopoulos, J. D. & Soljačić, M. Topological photonics. *Nat. Photon.* **8**, 821–829 (2014).
- Ozawa, T. et al. Topological photonics. *Rev. Mod. Phys.* **91**, 015006 (2019).
- Sundaram, G. & Niu, Q. Wave-packet dynamics in slowly perturbed crystals: gradient corrections and Berry-phase effects. *Phys. Rev. B* **59**, 14915–14925 (1999).
- Yang, Z. et al. Topological acoustics. *Phys. Rev. Lett.* **114**, 114301 (2015).
- Cooper, N. R., Dalibard, J. & Spielman, I. B. Topological bands for ultracold atoms. *Rev. Mod. Phys.* **91**, 015005 (2019).
- Delplace, P., Marston, J. & Venaille, A. Topological origin of equatorial waves. *Science* **358**, 1075–1077 (2017).
- Zanardi, P., Giorda, P. & Cozzini, M. Information-theoretic differential geometry of quantum phase transitions. *Phys. Rev. Lett.* **99**, 100603 (2007).
- Liang, L., Peotta, S., Harju, A. & Törmä, P. Wave-packet dynamics of Bogoliubov quasiparticles: quantum metric effects. *Phys. Rev. B* **96**, 064511 (2017).
- Hauke, P., Lewenstein, M. & Eckardt, A. Tomography of band insulators from quench dynamics. *Phys. Rev. Lett.* **113**, 045303 (2014).
- Lim, L.-K., Fuchs, J.-N. & Montambaux, G. Geometry of Bloch states probed by Stückelberg interferometry. *Phys. Rev. A* **92**, 063627 (2015).
- Fläschner, N. et al. Experimental reconstruction of the Berry curvature in a Floquet Bloch band. *Science* **352**, 1091–1094 (2016).
- Wimmer, M., Price, H. M., Carusotto, I. & Peschel, U. Experimental measurement of the Berry curvature from anomalous transport. *Nat. Phys.* **13**, 545–550 (2017).
- Yu, M. et al. Experimental measurement of the quantum geometric tensor using coupled qubits in diamond. *Natl Sci. Rev.* nww193 (2019).
- Tan, X. et al. Experimental measurement of the quantum metric tensor and related topological phase transition with a superconducting qubit. *Phys. Rev. Lett.* **122**, 210401 (2019).
- Bleu, O., Solnyshkov, D. D. & Malpuech, G. Measuring the quantum geometric tensor in two-dimensional photonic and exciton-polariton systems. *Phys. Rev. B* **97**, 195422 (2018).
- Richter, S. et al. Exceptional points in anisotropic planar microcavities. *Phys. Rev. A* **95**, 023836 (2017).
- Nalitov, A. V., Solnyshkov, D. D. & Malpuech, G. Polariton Z topological insulator. *Phys. Rev. Lett.* **114**, 116401 (2015).
- Klemmt, S. et al. Exciton-polariton topological insulator. *Nature* **562**, 552–556 (2018).
- St-Jean, P. et al. Lasing in topological edge states of a one-dimensional lattice. *Nat. Photon.* **11**, 651–656 (2017).
- Bleu, O., Malpuech, G. & Solnyshkov, D. D. Robust quantum valley Hall effect for vortices in an interacting bosonic quantum fluid. *Nature Commun.* **9**, 3991 (2018).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Methods

Sample details

The sample used for this experiment is a high quality-factor 3/2λ GaAs/AlGaAs planar cavity. The high growth precision of the sample via molecular beam epitaxy results in a sample with a quality factor exceeding 100,000, and the associated lifetime, τ , for the lower polariton branch surpasses 100 ps. This lifetime is measured by propagation experiments at negative exciton–photon detuning, as done previously^{34,35}. Moreover, the quantum wells show large areas (up to several hundreds of micrometres) that are free from defects. The cavity contains 12 GaAs quantum wells, 7-nm thick, organized in groups of four and placed at three antinode positions of the electric field. The front (back) mirror consists of 34 (40) pairs of AlAs/Al_{0.80}Ga_{0.20}As layers. The exciton energy is $E_{\text{exc}} = 1.611$ eV and the Rabi splitting is $\hbar\Omega_R \approx 16$ meV at 0 T. The full polariton dispersion measurement evidencing the exciton–photon anticrossing is shown in Supplementary Fig. 6.

The exciton diamagnetic shift (blueshift) is 4 meV at 9 T. The oscillator strength increases when applying 9 T, leading to a 17% increase in the Rabi splitting. These two effects almost cancel each other, resulting in a small blueshift for the bottom of the lower polariton branch of 0.4 meV at 9 T. The exciton–photon detuning (see below) at 9 T is therefore more negative by 4 meV than at 0 T. This detuning change remains moderate compared with the Rabi splitting value. The exciton fraction of the exciton polariton is reduced from 0.5 at 0 T to about 0.4 at 9 T, whereas the photon fraction increases from 0.5 to about 0.6. The exciton diamagnetic shift and the increase of the exciton oscillator strength are well known effects that are the result of the decrease of the exciton Bohr radius under a magnetic field^{36,37}.

Exciton–photon detuning

One of the most important parameters controlling the properties of exciton polaritons is the exciton–photon detuning¹⁰, which is the difference between the energies of the two bare resonances at $k = 0$: $\delta = E_{\text{phot}} - E_{\text{exc}}$. This parameter controls the excitonic and photonic fractions of the lower polariton branch—that is, their deviation from equal values of 50% at $\delta = 0$. For example, the excitonic fraction (the square of the Hopfield coefficient) for relatively small detunings is $x(\delta) \approx 1/2 + \delta/2\hbar\Omega_R$. As an example, as the detuning becomes more negative, the polaritons become more photonic, which means (among other things) that the exciton-related Zeeman splitting Δ_z decreases and the photon-related spin–orbit coupling β increases.

To perform the two experiments—the QGT and the anomalous Hall drift measurements—we selected two different regions of the same wafer. For the QGT measurements, we selected a central region of the sample at $\delta = 0$ meV and 0 T. Then, to have the adiabatic acceleration needed to observe the anomalous Hall drift effect, we selected a lateral region of the same wafer showing a rapid change in the energy of the lower polariton branch (approximately equal to 3 nm mm^{−1}), and a negative detuning of $\delta \approx -10$ meV.

Experimental setup to measure the QGT

The microcavity is cooled to 4 K in a closed-loop helium cryostat equipped with a superconductive magnet able to generate a field onto the sample that spans from −9 T to 9 T in a Faraday configuration (the external magnetic field is perpendicular to the microcavity plane).

For the measurement of the QGT, the excitation is performed by an off-resonance linearly polarized continuous-wave 2-μm-diameter laser spot, tuned to the first minimum of the stopband oscillations, so as to maximize the injection. The sample excitation and the collection of the polaritonic photoluminescence is performed in a reflection scheme, by means of a wide numerical aperture objective (0.86), resulting in a 14-μm^{−1} field of view. A 8-μm^{−1} portion of k space is then reconstructed

on the monochromator slits so that the photoluminescence can be energetically resolved. To avoid any loss of k -space information, the entire detection line is built in a 2f configuration and the required polarization filtering is performed in the real-space plane. The polarization response of the setup is characterized before the experiments. The raw photoluminescence data are collected by an automatic Labview routine able to perform a complete tomography in any of the three polarization bases (HV, AD or RL), via the sequential passage of light through a pair of motorized quarter- and half-waveplates and a polarizer. The energy mapping onto the charge-coupled device (CCD) camera is performed through a 550-cm monochromator equipped with a grating of 1,800 lines per mm and slit aperture set to 80 μm. For each polarization, a scan of 561 images is acquired, each containing an $I(E, k_v)$ map at a given k_H , upon moving a translational stage mounting the final lens by steps of 12 μm. In this way, a three-dimensional set of photoluminescent data, $I(E, k_H, k_v)$, is collected in any of the six polarization states. The energy resolution of the image is $\delta E = 0.015$ nm. The momentum resolutions are $\delta k_v = 0.008$ μm^{−1} per pixel and $\delta k_H = 0.014$ μm^{−1} per frame, corresponding to the momentum magnification with respect to the CCD pixel size and scan lens movement step, respectively.

Experimental setup to measure the polariton anomalous Hall effect

The anomalous Hall drift experiment is also performed in reflection configuration, in the same cryostat at the same temperature and magnetic field conditions, but using a 3-cm focal distance doublet ensuring a real-space field of view exceeding 500 μm. Resonant excitation of the lower polariton mode at $k = 0$ is performed, with a polarization of excitation which is adjusted to the one of the eigenstates. The chosen region of the sample exhibits the highest gradients³⁴.

The experimental uncertainties for all types of measurements are discussed in Supplementary Note 5.

Numerical analysis

We start by fitting the total intensity for each wavevector (k_H, k_v) with a double Gaussian curve, which enables the discovery of the energies of the two eigenstates, E_{\pm} . Then, the intensities of the six polarization components are obtained at the energies of the eigenstates by integration within the Gaussian width, and the components of the pseudospin calculated from these intensities. If the modes are almost degenerate in total intensity, with the energy difference falling below the inhomogeneous broadening, they can often still be distinguished by separately studying the spectra in the polarization components. This enables resolution of the branches for small energy differences. The pseudospin maps of the lower and upper eigenstates, encoded in the angles θ and ϕ , are then smoothed with a low-pass filter that eliminates noise. Finally, the components of the QGT are calculated from the pseudospin using equation (4). The gradient is obtained by the Green–Gauss method with simple face averaging. Parallel computing is used to accelerate the treatment of 4.6×10^9 experimental datapoints. The final resolution of the QGT components is $1,024 \times 561$.

Data availability

The datasets generated and/or analysed during the current study are available in the Open Science Framework (OSF) repository at https://osf.io/s4rzu/?view_only=1cabd49416c04a9baed856dee3ae1ba9.

34. Steger, M., Gautham, C., Snoke, D. W., Pfeiffer, L. & West, K. Slow reflection and two-photon generation of microcavity exciton–polaritons. *Optica* **2**, 1–5 (2015).

35. Ballarini, D. et al. Macroscopic two-dimensional polariton condensates. *Phys. Rev. Lett.* **118**, 215301 (2017).

36. Armitage, A. et al. Exciton polaritons in semiconductor quantum microcavities in a high magnetic field. *Phys. Rev. B* **55**, 16395–16403 (1997).

37. Rahimi-Iman, A. et al. Zeeman splitting and diamagnetic shift of spatially confined quantum-well exciton polaritons in an external magnetic field. *Phys. Rev. B* **84**, 165325 (2011).

Acknowledgements We thank D. Colas for critical reading of the manuscript. This work was supported by the ERC project ElecOpteR (grant number 780757). We acknowledge the support of the project Quantum Fluids of Light (ANR-16-CE30-0021), of the ANR Labex Ganex (ANR-11-LABX-0014), and of the ANR program Investissements d'Avenir through the IDEX-ISITE initiative 16-IDEX-0001 (CAP 20-25). D.D.S. acknowledges the support of the IUF (Institut Universitaire de France). This work was partially supported by the FISIR-CNR project "TECNOMED—Tecnopolo di nanotecnologia e fotonica per la medicina di precisione".

Author contributions A.G., L.D. and D.B. designed the setup. A.G. realized the experiments with the help of V.A., M.D.G. and G.L. D.S. supervised the experimental part. K.W.W. and L.N.P. fabricated the sample. O.B., D.D.S. and G.M. performed the treatment of the experimental data. O.B. performed analytical calculations. G.M. and O.B. wrote the manuscript with input from all authors.

Competing interests The authors declare no competing interests.

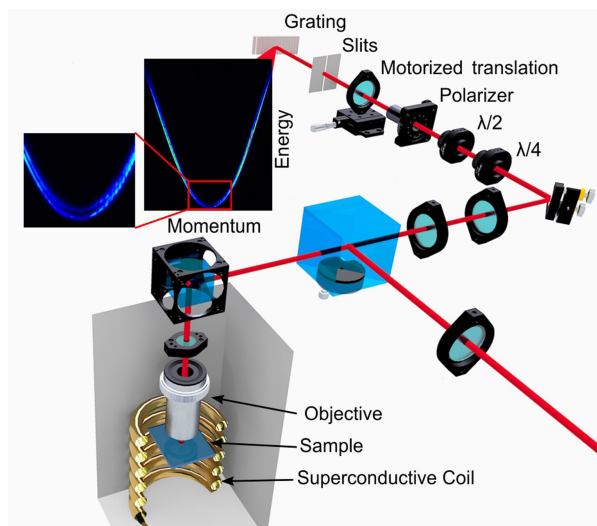
Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-1989-2>.

Correspondence and requests for materials should be addressed to D.S. or G.M.

Peer review information *Nature* thanks Ulf Peschel and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | Experimental setup. Schematic of the polarization tomography experiment. The incoming pump laser (bottom right) is focused onto the sample held in the cryogenic superconductive magnet (bottom left). The emission is recollimated, polarization filtered and the momentum space optically rebuilt at the entrance slits of a spectrometer (top) with energy resolution of $30 \mu\text{eV}$ (top left). The Zeeman splitting is highlighted in the inset.

Attosecond pulse shaping using a seeded free-electron laser

<https://doi.org/10.1038/s41586-020-2005-6>

Received: 24 May 2019

Accepted: 5 November 2019

Published online: 10 February 2020

 Check for updates

Praveen Kumar Maraju¹, Cesare Grazioli², Michele Di Fraia³, Matteo Moiola¹, Dominik Ertel¹, Hamed Ahmadi¹, Oksana Plekan³, Paola Finetti³, Enrico Allaria³, Luca Giannessi^{3,4}, Giovanni De Ninno^{3,5}, Carlo Spezzani³, Giuseppe Penco³, Simone Spampinati³, Alexander Demidovich³, Miltcho B. Danailov³, Roberto Borghes³, George Kourousias³, Carlos Eduardo Sanches Dos Reis³, Fulvio Billé³, Alberto A. Lutman⁶, Richard J. Squibb⁷, Raimund Feifel⁷, Paolo Carpeggiani⁸, Maurizio Reduzzi⁹, Tommaso Mazza¹⁰, Michael Meyer¹⁰, Samuel Bengtsson¹¹, Neven Ibrakovic¹¹, Emma Rose Simpson¹¹, Johan Mauritsson¹¹, Tamás Csizmadia¹², Mathieu Dumergue¹², Sergei Kühn¹², Harshitha Nandiga Gopalakrishna¹², Daehyun You¹³, Kiyoshi Ueda¹³, Marie Labeye¹⁴, Jens Egebjerg Bækthøj¹⁴, Kenneth J. Schafer¹⁴, Elena V. Gryzlova¹⁵, Alexei N. Grum-Grzhimailo¹⁵, Kevin C. Prince³, Carlo Callegari³ & Giuseppe Sansone^{1✉}

Attosecond pulses are central to the investigation of valence- and core-electron dynamics on their natural timescales^{1–3}. The reproducible generation and characterization of attosecond waveforms has been demonstrated so far only through the process of high-order harmonic generation^{4–7}. Several methods for shaping attosecond waveforms have been proposed, including the use of metallic filters^{8,9}, multilayer mirrors¹⁰ and manipulation of the driving field¹¹. However, none of these approaches allows the flexible manipulation of the temporal characteristics of the attosecond waveforms, and they suffer from the low conversion efficiency of the high-order harmonic generation process. Free-electron lasers, by contrast, deliver femtosecond, extreme-ultraviolet and X-ray pulses with energies ranging from tens of microjoules to a few millijoules^{12,13}. Recent experiments have shown that they can generate subfemtosecond spikes, but with temporal characteristics that change shot-to-shot^{14–16}. Here we report reproducible generation of high-energy (microjoule level) attosecond waveforms using a seeded free-electron laser¹⁷. We demonstrate amplitude and phase manipulation of the harmonic components of an attosecond pulse train in combination with an approach for its temporal reconstruction. The results presented here open the way to performing attosecond time-resolved experiments with free-electron lasers.

The intensities and relative phases between the harmonics $q\omega_f$ (with q an integer and ω_f the fundamental frequency) in an extreme ultraviolet (XUV) frequency comb determine the temporal structure of the resulting attosecond pulse train. The intensities of the harmonics can be easily measured using a (photon or electron) spectrometer. Phase information, which is harder to come by, is usually obtained by observing the interference between different pathways leading to states with the same final energy, where the phase to be characterized is included in at least one of the pathways. With XUV pulses, the natural observable is a photoelectron, hence different pathways into the ionization continuum are studied. The XUV frequency comb produced by high-order harmonic generation (HHG) consists of odd-integer harmonics of the fundamental field, and the ionization process takes

place in the presence of a near-infrared (NIR) dressing field with the same frequency ω_f . Under these conditions, additional photons may be absorbed or emitted, producing a single sideband halfway between the main photoelectron peaks. Each sideband can be populated through two pathways leading to final states of the same parity and this results in a variation in sideband amplitude as a function of the relative phase of the two pathways. If the XUV and fundamental fields are precisely synchronized, as they can be in HHG, then delaying the fields with respect to each other reveals the phase information^{4,18}.

In our study, the harmonic comb was generated by the seeded free-electron laser (FEL) FERMI, which uses an ultraviolet pulse ($\omega_{UV} = \omega_f = 4.69$ eV) derived from a frequency-tripled NIR pulse ($\omega_{NIR} = \omega_{UV}/3$) as the seed. Three ($q = 7, 8, 9$) and four ($q = 7, 8, 9, 10$)

¹Physikalisches Institut, Albert-Ludwigs-Universität, Freiburg, Germany. ²ISM-CNR, Trieste LD2 Unit, Trieste, Italy. ³Elettra-Sincrotrone Trieste SCpA, Basovizza, Trieste, Italy. ⁴Istituto Nazionale di Fisica Nucleare, Laboratori Nazionali di Frascati, Rome, Italy. ⁵Laboratory of Quantum Optics, University of Nova Gorica, Nova Gorica, Slovenia. ⁶SLAC National Accelerator Laboratory, Menlo Park, CA, USA. ⁷Department of Physics, University of Gothenburg, Gothenburg, Sweden. ⁸Institut für Photonik, Technische Universität Wien, Vienna, Austria. ⁹Dipartimento di Fisica, Politecnico di Milano, Milan, Italy. ¹⁰European XFEL GmbH, Schenefeld, Germany. ¹¹Department of Physics, Lund University, Lund, Sweden. ¹²ELI-ALPS, ELI-Hu Kft, Szeged, Hungary. ¹³Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, Sendai, Japan. ¹⁴Department of Physics and Astronomy, Louisiana State University, Baton Rouge, LA, USA. ¹⁵Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia. ✉e-mail: giuseppe.sansone@physik.uni-freiburg.de

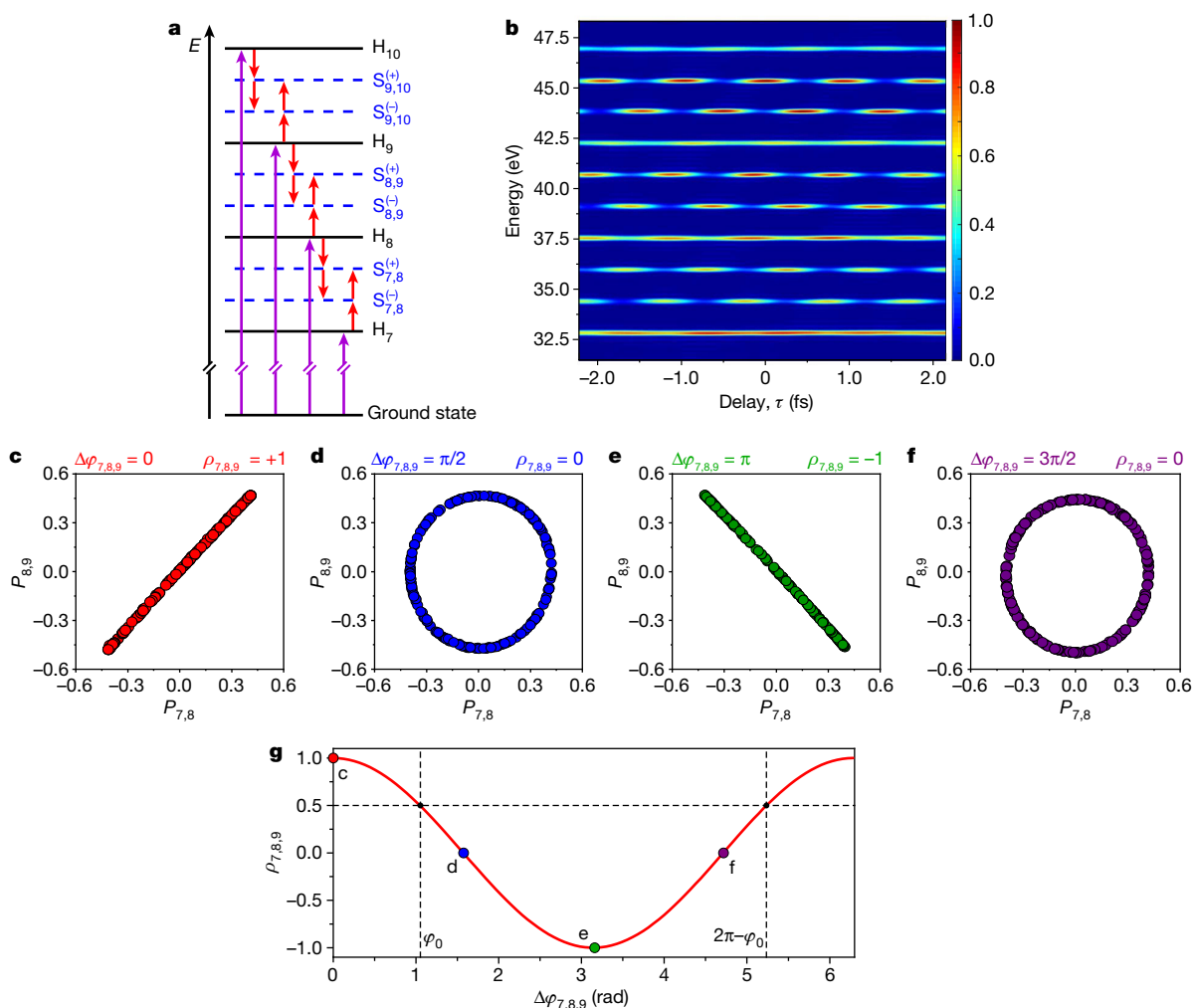


Fig. 1 | Multi-photon sideband generation and principle of the measurement. **a**, Schematic view of multi-NIR-photon sideband generation. Shown are the energy levels of the photoelectrons generated by the harmonics of the FEL (H_7 – H_{10} ; magenta arrows) and by the additional absorption and emission of one and two NIR photons ($S_{q,q+1}^{(\pm)}$; red arrows). **b**, Expected normalized intensity of the photoelectron spectra as a function of the relative delay τ between the train of attosecond pulses and the NIR field along the (positive) common direction of polarization of the two fields. The

photoelectron spectra are characterized by an oscillation with a period $T = 2\pi/(3\omega_{\text{NIR}})$. **c–f**, Correlation plots of the oscillating components of the sidebands ($P_{7,8}$, $P_{8,9}$) for four phase differences $\Delta\varphi_{7,8,9}$; 0 (**c**), $\pi/2$ (**d**), π (**e**) and $3\pi/2$ (**f**). At the top of each plot is shown the value of $\Delta\varphi_{7,8,9}$ (left) and $\rho_{7,8,9}$ (right). **g**, Evolution of the correlation coefficient $\rho_{7,8,9}$ as a function of $\Delta\varphi_{7,8,9}$, showing the locations corresponding to the plots **c–f**. The intensity of the NIR pulse is $I_{\text{NIR}} = 1.5 \times 10^{11} \text{ W cm}^{-2}$.

harmonics of ω_{UV} were generated using two different undulator configurations (see Extended Data Fig. 1 and Extended Data Table 1). To characterize the pulses, photoionization took place in the presence of a field with frequency ω_{NIR} leading to the formation of two sidebands between each pair of the main XUV peaks (see Fig. 1a). The two sidebands $S_{q,q+1}^{(\pm)}$ can be each populated through two paths characterized by a different number of exchanged NIR photons: the absorption of one photon of the harmonic q and one (two) NIR photon, or the absorption of one photon of the harmonic $q+1$ and the emission of two (one) NIR photons. The difference in parity of the final states of the interfering pathways determines an asymmetry in the photoionization emission. If the observation is restricted along a single direction, the intensity of the sidebands oscillates as a function of the delay τ between the NIR and XUV pulse (see Fig. 1b)

$$S_{q,q+1}^{(\pm)}(\tau) \propto 1 + \alpha_{q,q+1} \cos[\varphi_{q+1} - \varphi_q + 3\omega_{\text{NIR}}\tau] = 1 \pm P_{q,q+1}(\tau) \quad (1)$$

where $\alpha_{q,q+1}$ depends on the intensity and energy of the two harmonics q and $q+1$ with phases φ_q and φ_{q+1} , on the photoelectron energy, and

on the intensity of the NIR pulse, and the equality defines the oscillating component of the sideband intensity $P_{q,q+1}$ under the approximations detailed in the Supplementary Information. If the delay τ could be precisely controlled, then the relative phase between consecutive harmonics could be estimated from the time shift between the oscillations of the sidebands.

This approach cannot be applied directly to the reconstruction of the relative phase of multiple harmonics generated by an FEL owing to the lack of subcycle synchronization between the harmonics and the NIR field¹⁹, which completely washes out the delay dependence of the sideband oscillations. The information can be still retrieved, however, through a correlation analysis of the fluctuating sideband intensities measured on a single-shot basis. This approach is presented in Fig. 1c–f, which shows the simulated correlation plots of the oscillating components $P_{8,9}$ and $P_{7,8}$ of the sidebands for a random variation of the delay τ in the range ± 3 fs, which is the typical delay jitter measured in the experiment²⁰. The correlation plot eliminates the explicit dependence on τ and results in an ellipse, the shape of which depends on the phase difference:

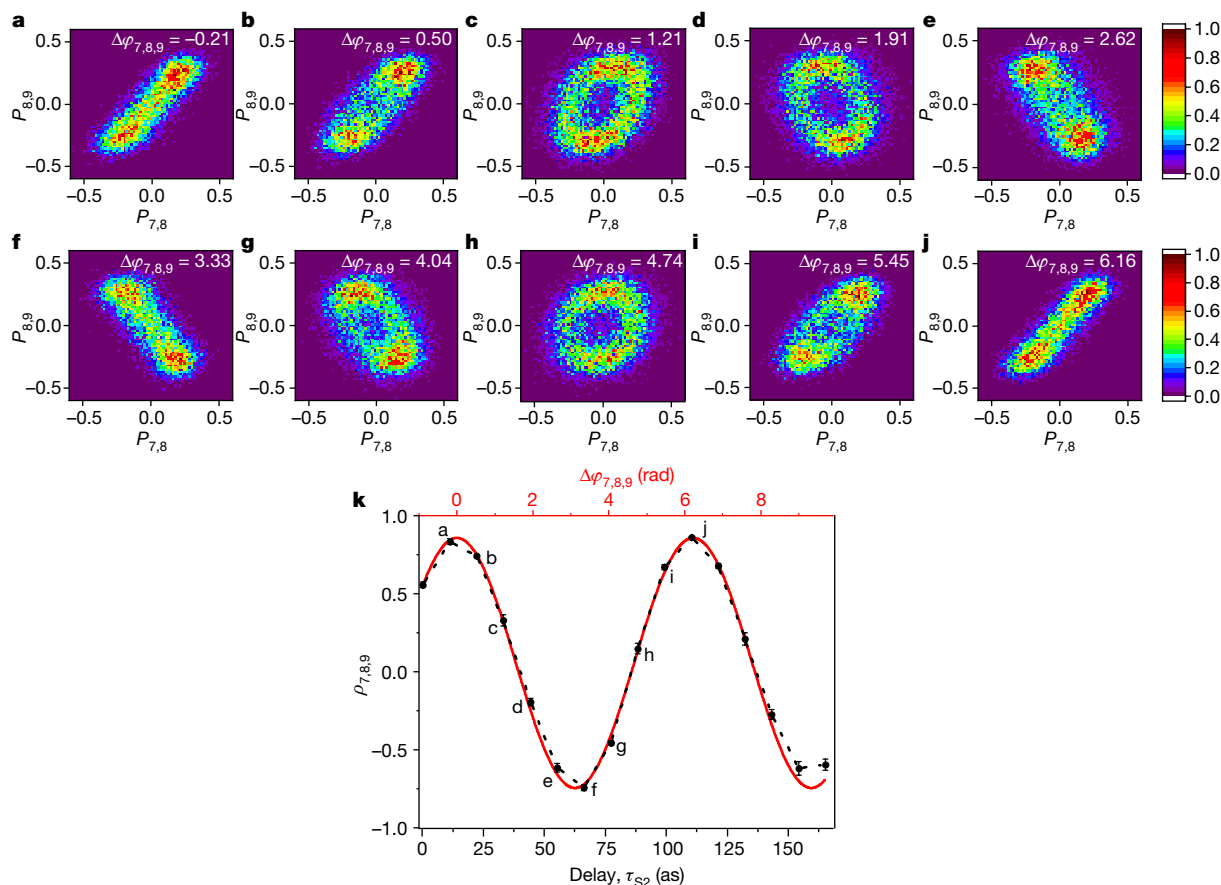


Fig. 2 | Correlation plots of the oscillating components of the sidebands and retrieval of phase difference $\Delta\phi_{7,8,9}$. **a–j**, Evolution of the correlation plots of the oscillating components of the sidebands ($P_{7,8}$, $P_{8,9}$) for increasing values of the delay τ_{s2} introduced by the phase shifter PS_2 (Extended Data Fig. 1a). At the top right of each panel is given the value of $\Delta\phi_{7,8,9}$. The colour scale indicates the density of single-shot experimental points normalized to unity for each panel. **k**, Evolution of the correlation coefficient $\rho_{7,8,9}$ as a function of the delay τ_{s2} (black points and dashed line) and the sinusoidal fit (red). The $\Delta\phi_{7,8,9}$ upper x

axis (red) was obtained by assigning the maxima of the fit to the values $\Delta\phi_{7,8,9} = 2m\pi$, where m is an integer. Letters a–j show the locations corresponding to the data shown in the panels above. The error bars are the standard deviation of the average correlation coefficient $\rho_{7,8,9}$ evaluated over ten sets of experimental data (each set consists of 1,200 single-shot points). The intensity of the NIR pulse was estimated to be $I_{\text{NIR}} \approx 1.5 \times 10^{12} \text{ W cm}^{-2}$. The value of the correlation parameters, the phase differences and the corresponding errors are presented in Extended Data Table 2.

$$\Delta\phi_{q-1,q,q+1} = \phi_{q+1} + \phi_{q-1} - 2\phi_q \quad (2)$$

The intensity profile of the pulse train depends only on this phase difference (apart from a trivial time shift; see Supplementary Information). Depending on the phase difference $\Delta\phi_{7,8,9}$, the plot evolves from a linear distribution with positive correlation (Fig. 1c), to a circle (Fig. 1d), to a linear distribution with negative correlation (Fig. 1e), and finally back to a circle (Fig. 1f). These changes clearly indicate that the shape of the correlated distribution is related to the synchronization of the three harmonics (the complete evolution as a function of the phase difference is presented in Extended Data Fig. 2). The phase information can be derived from the distribution by evaluating its correlation coefficient $\rho_{q-1,q,q+1}$ (see Supplementary Information), which quantifies the extent to which the two oscillating components oscillate perfectly in phase ($\Delta\phi_{q-1,q,q+1} = 0$ and $\rho_{q-1,q,q+1} = +1$) or out of phase ($\Delta\phi_{q-1,q,q+1} = \pi$ and $\rho_{q-1,q,q+1} = -1$). The correlation coefficient oscillates as a function of the phase difference $\Delta\phi_{7,8,9}$ as shown Fig. 1g, and it closely resembles a cosine function. Two different values of $\Delta\phi_{7,8,9}$ correspond to the same value of the correlation coefficient: ϕ_0 and $2\pi - \phi_0$. This ambiguity can be resolved in the experiment by controlling the modulus and sign of the phase differences between the harmonics (see Supplementary Information). Simulations based on the solution of the time-dependent Schrödinger equation confirmed the validity of our approach (see Extended Data Fig. 3).

In the experiment, the intensity of the harmonics was independently controlled by tuning the undulator gaps and the dispersive section of the electron transport optics. The phase between the harmonics was controlled by phase shifters^{21,22}, which introduce a delay τ_{si} (i indicates the i th-phase shifter) for a selected harmonic q , affecting the phase difference $\Delta\phi_{q-1,q,q+1}$ through a term $q\omega_{UV}\tau_{si}$ (see Extended Data Fig. 1). In this respect, seeded FELs offer a superior degree of control with respect to HHG sources, for which the intensities and phases of the single harmonic cannot be independently controlled.

Figure 2a–j presents the experimental results for the three-harmonic case, for different delays τ_{s2} introduced on the ninth harmonic. These measurements indicate a periodic evolution of the correlated distributions in close agreement with the theoretical prediction. A partial broadening of the distributions is attributed to the shot-to-shot fluctuations of the single harmonic intensity. The correlation coefficients $\rho_{7,8,9}$ (black points and dotted line in Fig. 2k) and the fit (red curve) clearly follow a cosine evolution in good agreement with the simulations. The maxima of the fit were assigned to the phase differences $\Delta\phi_{7,8,9} = 0, 2\pi$ (see upper x axis in Fig. 2k) and the curve was used to assign a phase difference $\Delta\phi_{7,8,9}$ to each delay τ_{s2} . The error in the estimation of the phase difference depends on the slope of the curve (which depends on the NIR intensity) and was typically in the range 0.05–0.1 rad for our experimental conditions (see Supplementary Information). The characterization of pulses with reproducible temporal structure gives

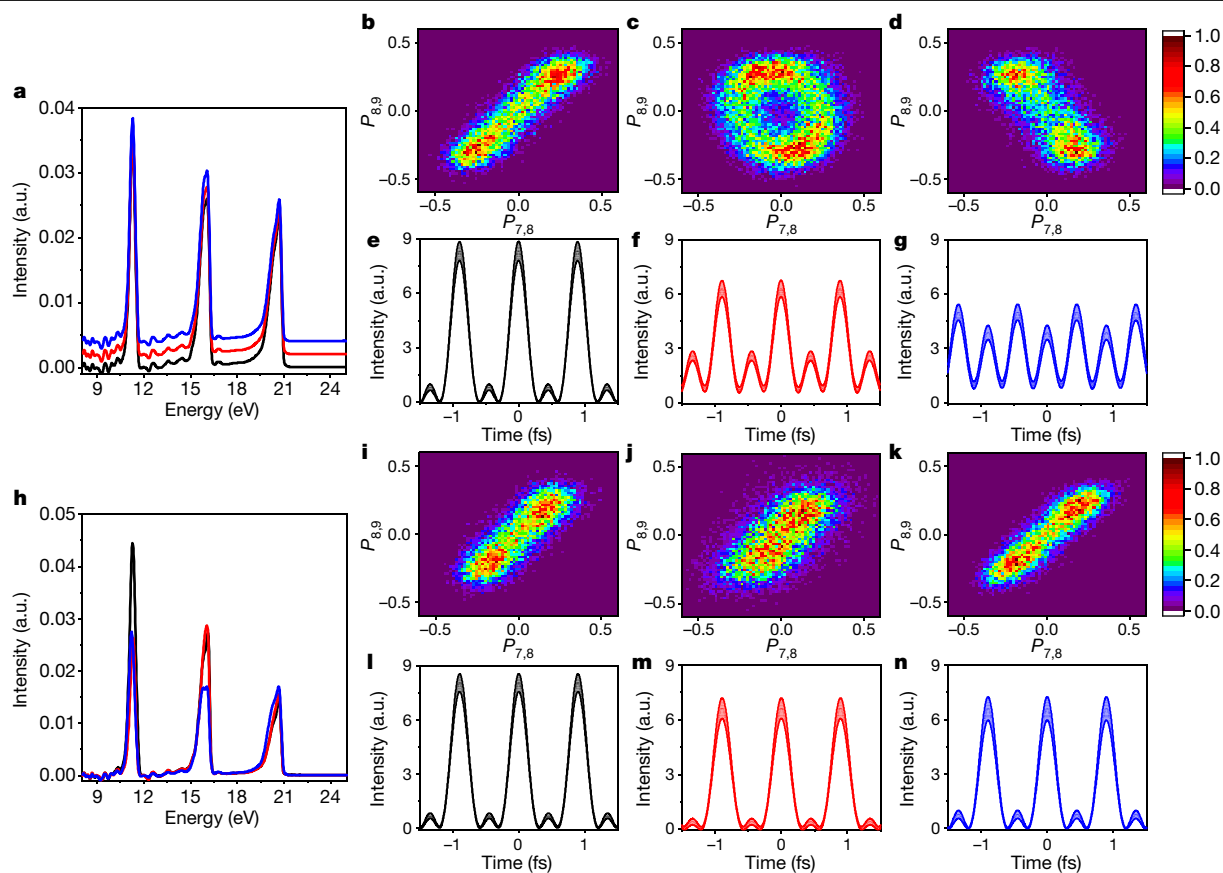


Fig. 3 | Complete phase and amplitude shaping of attosecond waveforms. **a–g**, Photoelectron spectra (**a**), correlation plots of the oscillating components of the sidebands ($P_{7,8}$, $P_{8,9}$; **b, c, d**) and reconstructed attosecond waveforms (**e, f, g**) in the case of independent phase shaping for three phase differences $\Delta\varphi_{7,8,9}$ (black curve and panels **b, e**; red curve and panels **c, f**; blue curve and panels **d, g**). The colour scale indicates the density of single-shot experimental points normalized to unity for each panel. The three photoelectron spectra in **a** were vertically shifted for visual clarity. **h–n**, Photoelectron spectra (**h**), correlation plots of the oscillating components of the sidebands ($P_{7,8}$, $P_{8,9}$; **i, j, k**) and reconstructed attosecond waveforms (**l, m, n**) in the case of independent

amplitude control for three settings of the harmonic amplitudes (black curve and panels **i, l**; red curve and panels **j, m**; blue curve and panels **k, n**) using the same values of the phase shifters. The colour scale indicates the density of single-shot experimental points normalized to unity for each panel. See Extended Data Table 3 for additional information on the phase difference $\Delta\varphi_{7,8,9}$ and the amplitude of the harmonics F_7 , F_8 and F_9 . The errors in the reconstruction of the attosecond pulse trains are determined by the error bars for the amplitudes and phase differences (see Extended Data Table 3) and are indicated as shaded areas in **e, f, g, l, m, n**.

the possibility of accumulating data over several single-shot measurements, thus improving the signal-to-noise ratio and reducing the error in the temporal reconstruction. In the case of a self-amplified spontaneous emission FEL, pulse properties change on a shot-to-shot basis and a single-shot technique is therefore mandatory¹⁶.

We exploited this approach to the determination of the relative phase of XUV harmonics to demonstrate the independent phase–amplitude shaping capability of attosecond waveforms offered by the FEL FERMI. Figure 3a and Fig. 3b, c and d show three photoelectron spectra and the corresponding correlation plots for three phase differences $\Delta\varphi_{7,8,9}$, respectively. The phase change does not appreciably modify the intensities of the three harmonics (for the amplitudes F_j of the single j th harmonic and for the phase differences see Extended Data Table 3). The reconstructed intensity profiles $I(t)$ are presented in Fig. 3e, f and g. The measurements indicate a pure phase shaping of the harmonic comb: a well-defined attosecond pulse train (Fig. 3e) obtained for $\Delta\varphi_{7,8,9} = 0.08 \pm 0.08$ rad (close to the ideal condition of harmonics in phase, $\Delta\varphi_{7,8,9} = 0$) is transformed first into an attosecond pulse train of lower amplitude with a satellite (Fig. 3f) when $\Delta\varphi_{7,8,9} = 1.93 \pm 0.03$ rad, and finally into an attosecond pulse train characterized by a double structure for $\Delta\varphi_{7,8,9} = 3.29 \pm 0.24$ rad (Fig. 3g), which is close to the condition of harmonics out of phase, $\Delta\varphi_{7,8,9} = \pi$. Figure 3h shows three photoelectron spectra corresponding to three different settings of the

amplitudes of the three harmonics (see Extended Data Table 3). The amplitude of the single harmonic was modified by about 25% using the dispersive section and the undulator gaps (see Supplementary Information). Figure 3i, j and k show the correlation plots for the same position of the phase shifters. The phase difference $\Delta\varphi_{7,8,9}$ remains constant within the experimental error, independent of the variations of the single harmonic intensity. The reconstructed attosecond pulse trains for the three configurations are presented in Fig. 3l, m and n. These data demonstrate a pure amplitude shaping of the harmonic comb: the well-defined attosecond pulse structure ($\Delta\varphi_{7,8,9}$ is close to zero (2π) for the three measurements) is preserved for the three configurations and the different harmonic intensities lead only to a variation in the maxima of the intensity profiles. The energy of the attosecond pulse train was about 16 μ J. Table-top-based HHG sources yield much lower energies (in the nanojoule range), and only a few experimental groups have reported total pulse energies on target approaching the microjoule range^{23,24}.

We estimated that a pulse energy of about 50 nJ per harmonic is sufficient for the acquisition of single-shot photoelectron spectra, which is well below the typical energy per harmonic (a few microjoules) available at FERMI. The currently available range of seed wavelengths at FERMI (360–230 nm) would allow a moderate control of the comb periodicity around 1 fs, but a revised layout of the seed laser optimized for this task

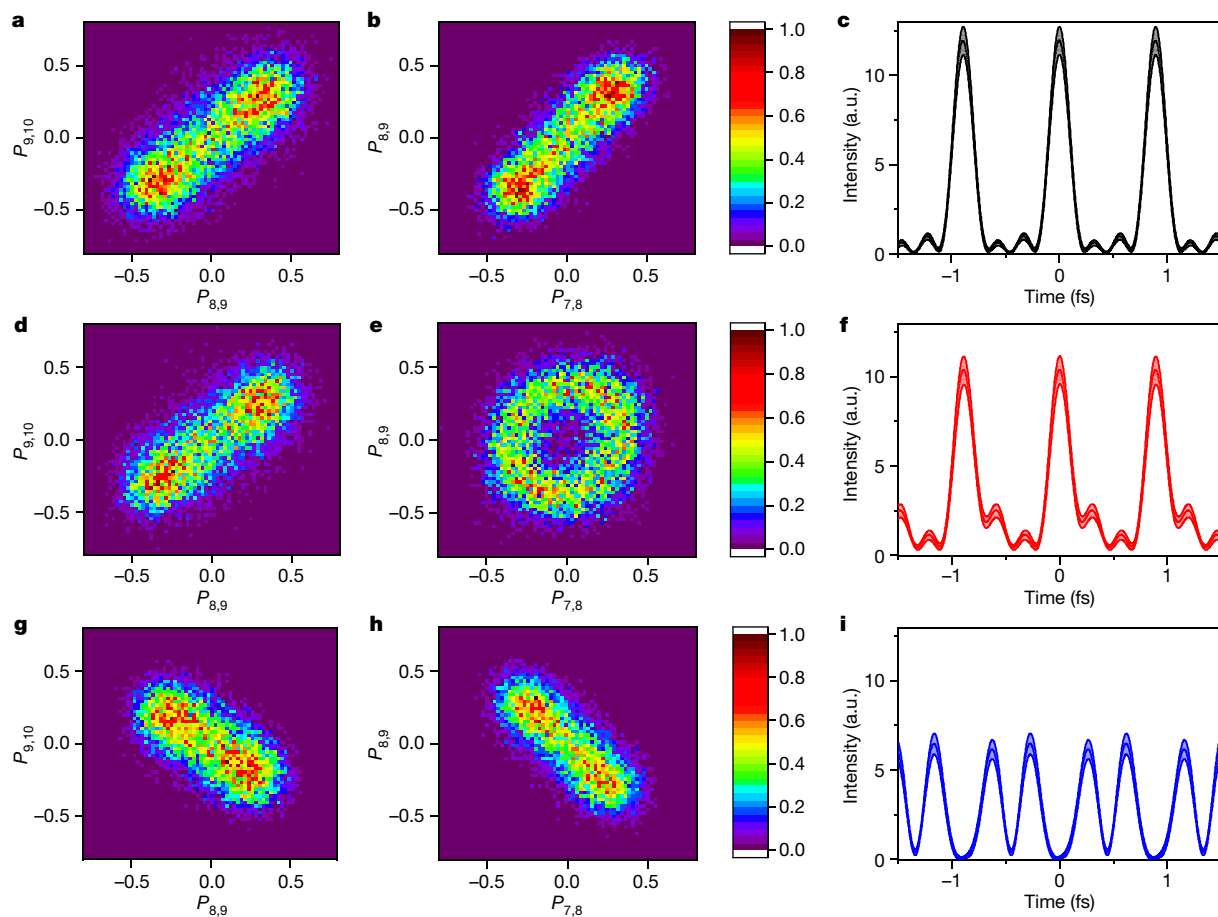


Fig. 4 | Synthesis of complex attosecond waveforms. a–i. Correlation plots of the oscillating components of the sidebands $P_{8,9}$, $P_{9,10}$ (**a, d, g**) and $P_{7,8}$, $P_{8,9}$ (**b, e, h**) and retrieved attosecond waveforms (**c, f, i**) for the four-harmonic experiment and three different combinations of the phase differences $\Delta\varphi_{7,8,9}$ and $\Delta\varphi_{8,9,10}$ (**a–c; d–f; g–i**). The colour scale indicates the density of single-shot experimental points normalized to unity for each panel. See Extended Data Table 3 for

additional information on the phase differences $\Delta\varphi_{7,8,9}$ and $\Delta\varphi_{8,9,10}$, and the amplitude of the harmonics F_7 , F_8 , F_9 and F_{10} . The errors in the reconstruction of the attosecond pulse trains are determined by the error bars for the amplitudes and phase differences (see Extended Data Table 3) and are indicated as shaded areas in **c, f, i**.

could increase the spike separation to tens of femtoseconds. Shorter separations can already be achieved by using only odd or only even harmonics (for example, $q = 6, 8, 10$).

We should point out that alternative FEL-based approaches have been theoretically proposed for the generation of a train of attosecond pulses^{25,26}. Even though the predicted peak power levels (gigawatts) and pulse durations (down to sub-100 as) are comparable with those reported here, or in principle achievable with our approach, these methods do not offer a strategy for controlling the relative amplitudes and phases of the single harmonics, that is, for attosecond pulse shaping. Extension of our approach to wavelengths as short as 4 nm (300 eV) appear feasible if combined with the echo-enabled harmonic generation seeding scheme²⁷. Numerical simulations indicate that the method for the temporal characterization could be applied for photon energies from about 20 eV up to 300 eV, using a suitable gas target.

As a first demonstration of complex attosecond waveform shaping, we considered the case of four harmonics (see Fig. 4), for which the attosecond temporal structure depends on the two phase differences $\Delta\varphi_{7,8,9}$ and $\Delta\varphi_{8,9,10}$. The photoelectron spectra with (red curve) and without (black curve) NIR are shown in Extended Data Fig. 1d. The independent control of the two phases gives the opportunity to generate ultrashort (chirp-free) attosecond pulse trains, as shown in Fig. 4a–c, which report the correlation plots ($P_{8,9}$ – $P_{9,10}$ in Fig. 4a and $P_{7,8}$ – $P_{8,9}$ in Fig. 4b) in the case of maximum positive correlation (see Extended Data Table 3 for the values of the amplitude and phase differences). The

reconstruction (Fig. 4c) returns a duration (FWHM) of the single pulse of about 210 ± 4 as. Figure 4d and e present the results corresponding respectively to $\Delta\varphi_{8,9,10} = 0.20 \pm 0.15$ rad (close to the configuration of harmonics in phase, $\Delta\varphi_{8,9,10} = 0$) and $\Delta\varphi_{7,8,9} = 1.23 \pm 0.06$ rad (harmonics only partially in phase). In the temporal domain, this condition translates into a partial broadening of the peaks (FWHM = 220 ± 5 as) and the appearance of small satellites in the reconstructed attosecond pulse train (Fig. 4f). Finally, Fig. 4g and h present the results when $\Delta\varphi_{8,9,10} = 2.89 \pm 0.08$ rad and $\Delta\varphi_{7,8,9} = 2.95 \pm 0.09$ rad (that is, both phase differences are close to π), respectively: the four harmonics are divided into two groups (harmonics 7 and 8, and harmonics 9 and 10), each pair of which is (approximately) in phase, with an additional phase jump of π between the two groups. This condition leads to a double attosecond pulse structure, which is visible in the reconstruction presented in Fig. 4i. The availability of six undulators at FERMI would in principle allow for the generation of six harmonics. This configuration may, however, require a revised and optimized setup. Simulations indicate that the experimental technique demonstrated in this work could be extended to the characterization of pulses with durations in the sub-100-as regime.

Our technique also offers the possibility of determining with subcycle resolution the relative phase of the XUV and NIR pulses, enabling phase-resolved pump–probe experiments at FELs based on attosecond pulse trains (see Supplementary Information and Extended Data Fig. 4). The high intensities in the XUV and X-ray spectral region reached by FELs,

combined with the capabilities offered by seeding to independently control and shape the amplitudes and phases of attosecond pulses, will open new possibilities for the investigation and control of ultrafast nonlinear electronic processes. The design of future seeded FEL sources at other facilities such as LCLS²⁸, FLASH²⁹ and SINAP³⁰ could be modified to optimize them for the mode of operation described here, which is at present possible at FERMI¹⁷ and DALIAN³¹. In solid samples, attosecond shaped waveforms could be used to promote electrons from the inner-valence to the conduction band, giving the possibility of investigating diffusion and relaxation effects with attosecond resolution and with temporally sculpted electronic wave packets. More generally, our results give unprecedented access to programmable attosecond waveforms at high intensities.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2005-6>.

- Krausz, F. & Ivanov, M. Attosecond physics. *Rev. Mod. Phys.* **81**, 163–234 (2009).
- Corkum, P. B. & Krausz, F. Attosecond science. *Nat. Phys.* **3**, 381–387 (2007).
- Kapteyn, H., Cohen, O., Christov, I. & Murnane, M. Harnessing attosecond science in the quest for coherent X-rays. *Science* **317**, 775–778 (2007).
- Paul, P. M. et al. Observation of a train of attosecond pulses from high harmonic generation. *Science* **292**, 1689–1692 (2001).
- Kienberger, R. et al. Atomic transient recorder. *Nature* **427**, 817–821 (2004).
- Tzallas, P., Charalambidis, D., Papadogiannis, N. A., Witte, K. & Tsakiris, G. D. Direct observation of attosecond light bunching. *Nature* **426**, 267–271 (2003).
- Nabekawa, Y. et al. Interferometric autocorrelation of an attosecond pulse train in the single-cycle regime. *Phys. Rev. Lett.* **97**, 153904 (2006).
- López-Martens, R. et al. Amplitude and phase control of attosecond light pulses. *Phys. Rev. Lett.* **94**, 033001 (2005).
- Gustafsson, E. et al. Broadband attosecond pulse shaping. *Opt. Lett.* **32**, 1353–1355 (2007).
- Hofstetter, M. et al. Attosecond dispersion control by extreme ultraviolet multilayer mirrors. *Opt. Express* **19**, 1767–1776 (2011).
- Bartels, R. et al. Shaped-pulse optimization of coherent emission of high-harmonic soft X-rays. *Nature* **406**, 164–166 (2000).
- Ackermann, W. et al. Operation of a free-electron laser from the extreme ultraviolet to the water window. *Nat. Photon.* **1**, 336–342 (2007).
- Emma, P. et al. First lasing and operation of an ångström-wavelength free-electron laser. *Nat. Photon.* **4**, 641–647 (2010).
- Marinelli, A. et al. Experimental demonstration of a single-spike hard-X-ray free electron-laser starting from noise. *Appl. Phys. Lett.* **111**, 151101 (2017).
- Huang, S. et al. Generating single-spike hard X-ray pulses with nonlinear bunch compression in free-electron lasers. *Phys. Rev. Lett.* **119**, 154801 (2017).
- Hartmann, N. et al. Attosecond time–energy structure of X-ray free-electron laser pulses. *Nat. Photon.* **12**, 215–220 (2018).
- Allaria, E. et al. Highly coherent and stable pulses from the FERMI seeded free-electron laser in the extreme ultraviolet. *Nat. Photon.* **6**, 699–704 (2012).
- Mairesse, Y. et al. Attosecond synchronization of high-harmonic soft X-rays. *Science* **302**, 1540–1543 (2003).
- Schulz, S. et al. Femtosecond all-optical synchronisation of an X-ray free-electron laser. *Nat. Commun.* **6**, 5938 (2015).
- Danailov, M. B. et al. Towards jitter-free pump-probe measurements at seeded free electron laser facilities. *Opt. Express* **22**, 12869–12879 (2014).
- Prince, K. C. et al. Coherent control with a short-wavelength free-electron laser. *Nat. Photon.* **10**, 176–179 (2016).
- lablonskyi, D. et al. Observation and control of laser-enabled Auger decay. *Phys. Rev. Lett.* **119**, 073203 (2017).
- Takahashi, E. J. et al. Attosecond nonlinear optics using gigawatt-scale isolated attosecond pulses. *Nat. Commun.* **4**, 2691 (2013).
- Nayak, A. et al. Multiple ionization of argon via multi-XUV-photon absorption induced by 20-GW high-order harmonic laser pulses. *Phys. Rev. A* **98**, 023426 (2018).
- Zholents, A. A. Method of an enhanced self-amplified spontaneous emission for x-ray free electron lasers. *Phys. Rev. Spec. Top. Accel. Beams* **8**, 040701 (2005).
- Thompson, N. R. & McNeil, B. W. J. Mode locking in a free-electron laser amplifier. *Phys. Rev. Lett.* **100**, 203901 (2008).
- Ribić, P. R. et al. Coherent soft X-ray pulses from an echo-enabled harmonic generation free-electron laser. *Nat. Photon.* **13**, 555–561 (2019).
- Hemsing, E. et al. *Soft X-ray FEL Seeding Studies for LCLS-II: Task Force Status Report. A White Paper by SLAC and LBNL*. Technical Note SLAC-TN-19-001 (SLAC, 2019); available at <https://www.slac.stanford.edu/pubs/slactns/tn06/slac-tn-19-001.pdf> (2019).
- Grattoni, V. et al. Status of seeding development at sFLASH. In *Proc. FEL2017* (eds Bishofberger, K., Carlsten, B. & Schaa, V. R. W.) 136–139 (JACoW, 2018).
- Zhao, Z. et al. Status of the SXFEL facility. *Appl. Sci.* **7**, 607 (2017).
- Yong, Y. et al. Dalian extreme ultraviolet coherent light source. *Chin. J. Lasers* **46**, 0100005 (2019).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Experimental setup

The experiment was performed at the seeded FEL FERMI and is schematically presented in Extended Data Fig. 1. Two different configurations of the undulators were implemented for the generation of three (Extended Data Fig. 1a) and four (Extended Data Fig. 1b) harmonics. The parameters of the FEL harmonics are reported in Extended Data Table 1. The seeding parameters (seed laser power and strength of the dispersive section) were carefully optimized in order to produce a sufficiently high bunching that could be preserved along the whole set of undulators tuned at the various harmonics. Tuning the first undulators to higher harmonics (shorter wavelengths) and the later ones to the lower harmonics had a twofold motivation. First, the bunching at higher harmonics was more prone to degradation and it would have been more difficult to preserve it up to the end of the undulator chain. Second, the diffraction at longer wavelengths was larger thus a shorter propagation path was preferable. For a properly optimized setup each undulator group produced a coherent, ~50-fs-long FEL pulse centred at the resonant wavelength³². The phase between the electric field of each harmonic was determined and controlled by the phase shifter available at FERMI at each undulator break.

The XUV and NIR pulses (energy $E_{\text{NIR}} = 45 \mu\text{J}$, duration $\text{FWHM}_{\text{NIR}} = 60 \text{ fs}$, intensity $I_{\text{NIR}} = 1.5 \times 10^{12} \text{ W cm}^{-2}$) were temporally and spatially overlapped in the interaction region with a residual shot-to-shot delay jitter of $\Delta\tau = \pm 3 \text{ fs}$ using a recombination mirror for collinear propagation. The single-shot photoelectron spectra, with (red lines) and without (black lines) the NIR pulse for the three and four harmonics configurations (see Extended Data Fig. 1c, d, respectively), were measured in neon using a magnetic bottle electron spectrometer placed in the interaction region (see Extended Data Fig. 1e). Only photoelectrons emitted in the upward hemisphere were collected, as shown by the detection efficiency (see Extended Data Fig. 1f) as a function of the angle between the emission direction and the spectrometer axis (which coincides with the (vertical) direction of polarization of the FEL pulses (see Supplementary Information)). For each machine setting, single-shot harmonic spectra (without NIR field) were measured³³. From these data we estimated a typical shot-to-shot fluctuation (standard deviation) for the intensity of each harmonic of about 5%–8%. The energy of the single harmonic was proportional to the integral of the corresponding peak of the photoelectron spectrum. The integral was corrected for the response function of the magnetic bottle spectrometer for different photoelectron energies, the cross-section of the target gas, and the transmission of the XUV beamline¹⁷.

The data were accumulated typically for about 10,000–12,000 shots for each machine and phase setting. For each setting, the mean intensities of the main photoelectron peaks I_q ($q = (7, 8, 9)$ or $q = (7, 8, 9, 10)$) were determined.

The total energy of the FEL pulse was measured (on a single-shot basis) with an ionization monitor placed upstream of the transmission and focusing XUV beamline.

Reconstruction of attosecond pulses

The temporal reconstruction of the attosecond pulse train using the correlation parameter $\rho_{q-1,q,q+1}$ for the simulations presented in Fig. 1c–f is shown in Extended Data Fig. 5a–d, respectively. The agreement between the input data (black curves) and the reconstructed profiles (red (Extended Data Fig. 5a), blue (Extended Data Fig. 5b), green (Extended Data Fig. 5c), and magenta (Extended Data Fig. 5d)) indicates the validity of our reconstruction method based on the value of the correlation parameter $\rho_{7,8,9}$. We also performed time-dependent Schrödinger equation (TDSE) simulations which confirmed the validity of our reconstruction protocol. The correlation curve obtained using the TDSE simulations reproduces that obtained by the strong field approximation (SFA) with a small shift of 0.157 rad, which results in only

minor corrections in the reconstruction of the intensity profile of the attosecond pulses. This shift was taken into account in the reconstructions presented in the manuscript.

The validity of the temporal reconstruction of the attosecond pulse train from the shift of the oscillations of the sidebands $S_{q-1,q}^{(+)} (S_{q-1,q}^{(-)})$ and $S_{q,q+1}^{(+)} (S_{q,q+1}^{(-)})$ is shown in Extended Data Fig. 5e, which reports the input (black lines) and reconstructed (blue dotted lines) intensity profiles for the simulation shown in Fig. 1b.

Simulations for attosecond pulse generation

The emission process for the configuration emitting harmonics 7, 8 and 9 can be simulated with the new version of the FEL code GENESIS 1.3³⁴. With the undulator set in the condition reported in Extended Data Fig. 1a, the resonant wavelength to be followed by the FEL code is different in the three sets of undulators. Given the fact that GENESIS 1.3 only tracks a relatively narrow band field this requires that different simulations are performed for the different sets of undulators. This option is normally used for harmonic generation FEL schemes and has been largely used in the study of the high-gain harmonic-generation operation mode implemented at FERMI. The problem is here complicated by the fact that consecutive undulators are tuned to wavelengths that are not integer multiples of one another. The recent upgrade of GENESIS 1.3 allows tracking of each single electron in the beam. The simulation can be performed if one carefully manages the transition from one undulator set to the next. After the interaction with the external seed laser is simulated, and the energy modulation at the 260 nm wavelength is imprinted in the beam, the particle phase-space is used to simulate the emission process at the ninth harmonic in the first group of undulators. The electric field produced here is then propagated in free space to the exit of the whole radiator, while the electrons are used for simulating the emission process at the eighth harmonic in the second group of undulators. The same is done with fields and electrons entering the third group of undulators. Finally, three fields are produced representing the emission from each undulator set at the exit of the radiator. The result of the simulation is presented in Extended Data Fig. 6a, b that shows the femtosecond envelope (Extended Data Fig. 6a) and the attosecond structure (Extended Data Fig. 6b) of the XUV pulse obtained using the combination of the three harmonics. The resonance condition, phase shifts between sets of undulators, and other parameters can be adjusted as usual. Simulations rely on the standard electron beam and seed laser for FERMI. Seed laser power and strength of the dispersive section are used as an optimization parameter to maximize the bunching and keep emission balanced between various harmonics. The parameters used in the simulation are summarized in Extended Data Table 4.

Data availability

Raw data were generated at the FERMI large-scale facility. Derived data supporting the findings of this study are available from the corresponding author on reasonable request.

32. Finetti, P. et al. Pulse duration of seeded free-electron lasers. *Phys. Rev. X* **7**, 021043 (2017).
33. Zangrando, M. et al. Recent results of PADReS, the Photon Analysis Delivery and REDuction System, from the FERMI FEL commissioning and user operations. *J. Synchrotron Rad.* **22**, 565–570 (2015).
34. Reiche, S. Update on the FEL code GENESIS 1.3. In *Proc. 36th Int. Free Electron Laser Conf. (FEL'14)* (eds Chrin, J., Reiche, S. & Schaa, V. R. W.) TUP019, 403–407 (IACoW, 2014).

Acknowledgements This project received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 641789 MEDEA and the Italian Ministry of Research (Project FIRB no. RBID08CRXX). K.U. acknowledges support from the X-ray Free Electron Laser Utilization Research Project and the X-ray Free Electron Laser Priority Strategy Program of the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT), from the Cooperative Research Program 'Network Joint Research Center for Materials and Devices: Dynamic Alliance for Open Innovation Bridging Human, Environment and Materials', from the bilateral project CNR-JSPS

'Ultrafast science with extreme ultraviolet Free Electron Lasers', and from the IMRAM project for international co-operation. R.F. and J.M. thank the Swedish Research Council (VR) and the Knut and Alice Wallenberg Foundation for financial support. E.V.G. acknowledges support from the Foundation for the Advancement of Theoretical Physics and Mathematics 'BASIS'. M. Meyer and T.M. acknowledge support from the Deutsche Forschungsgemeinschaft (DFG) under grant no. SFB925/1. A.A.L. was supported by the US Department of Energy contract no. DE-AC02-76SF00515. Research at Louisiana State University was supported by the US Department of Energy, Office of Science, Basic Energy Sciences, under contract no. de-sc0010431. Portions of this research were conducted with high performance computing resources provided by Louisiana State University (<http://www.hpc.lsu.edu>) and by Louisiana Optical Network Infrastructure (<http://hpc.loni.org>). G.S. acknowledges useful discussions about the simulations and the data analysis with T. Pfeifer and M. Kübel. We acknowledge L. Foglia, A. Simoncig and M. Coreno for valuable discussions.

Author contributions P.K.M., M. Moiola, D.E., M.D.F., O.P., H.A., P.C., T.M., M. Meyer, S.B., N.I., E.R.S., J.M., T.C., M.D., S.K., H.N.G., D.Y., K.U., K.C.P., C.G. and C.C. contributed to data acquisition and to data analysis. L.G., E.A., G.D.N., C.S., G.P. and S.S. operated the machine and designed the three and four harmonic generation scheme. A.A.L. contributed to the machine operation. A.D. and M. B. D. designed the beam path for the NIR pulse. P.F. designed the mechanics for the recombination mirror. A.D. and C.G. designed the recombination mirror, and the whole set-up was installed by A.D., C.G. and M.D.F. A.D. and

M.D.F. designed and installed the beam dump diagnostic system for alignment of the collinear configuration. M.D.F., O.P. and C.C. prepared the end station. R.B., G.K., C.E.S.D.R. and F.B. developed the analysis tools used during beamtime. M.R. contributed to the preliminary development of the simulation codes. M.L., J.E.B. and K.J.S. performed the TDSE calculations. R.J.S. and R.F. constructed and operated the magnetic bottle electron spectrometer. A.N.G.-G. and E.V.G. developed the perturbation-theory approach and derived the atomic phase contributions. C.C. and G.S. conceived the idea of the experiment. G.S. developed the numerical code for the SFA simulations. P.K.M. developed the numerical code for the correlation analysis. P.K.M., C.C. and G.S. analysed the experimental data and performed the simulations. G.S. supervised the work. P.K.M., A.N.G.-G., E.V.G., C.C. and G.S. wrote the manuscript, which was discussed and agreed by all coauthors.

Competing interests The authors declare no competing interests.

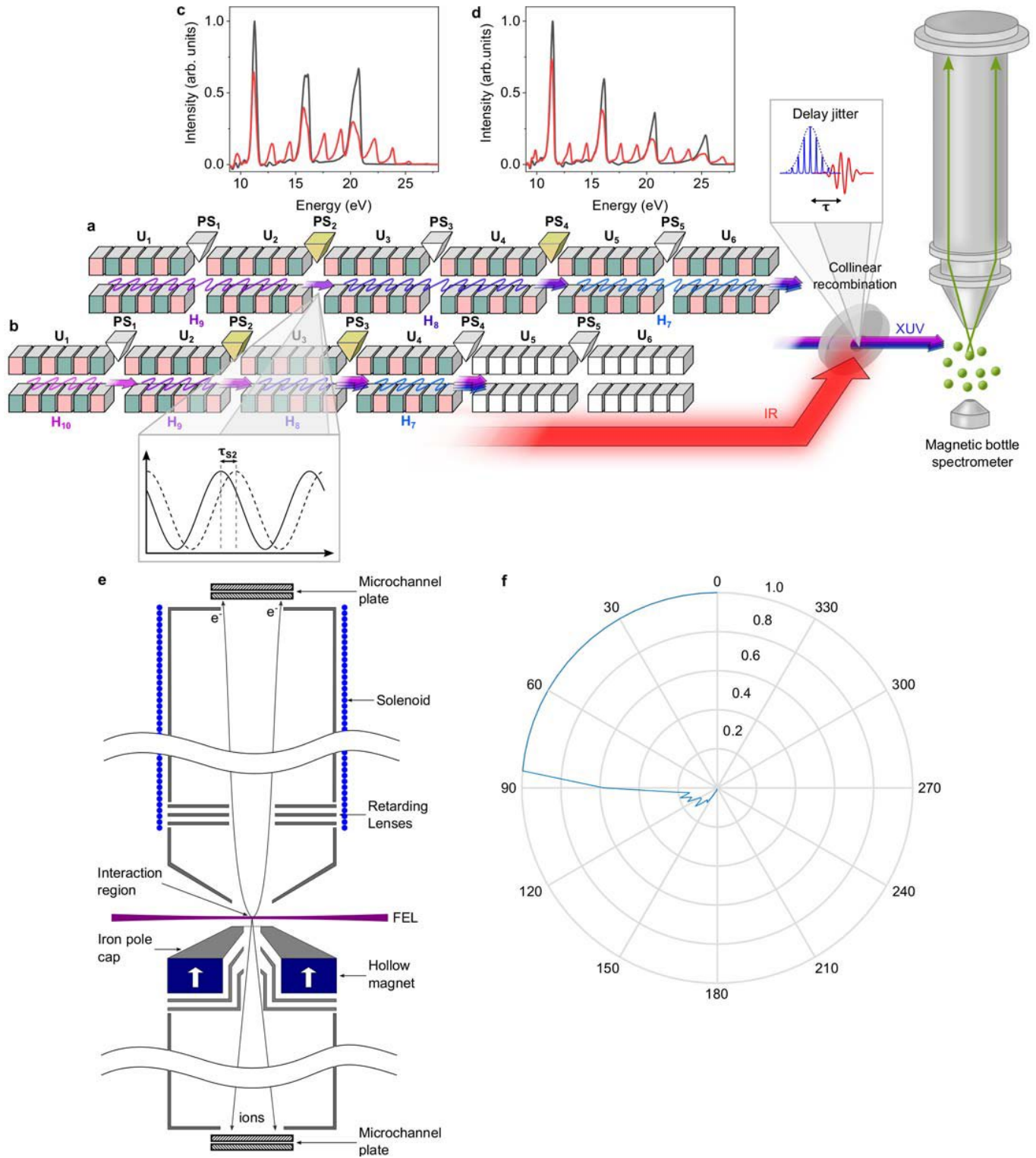
Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-2005-6>.

Correspondence and requests for materials should be addressed to G.S.

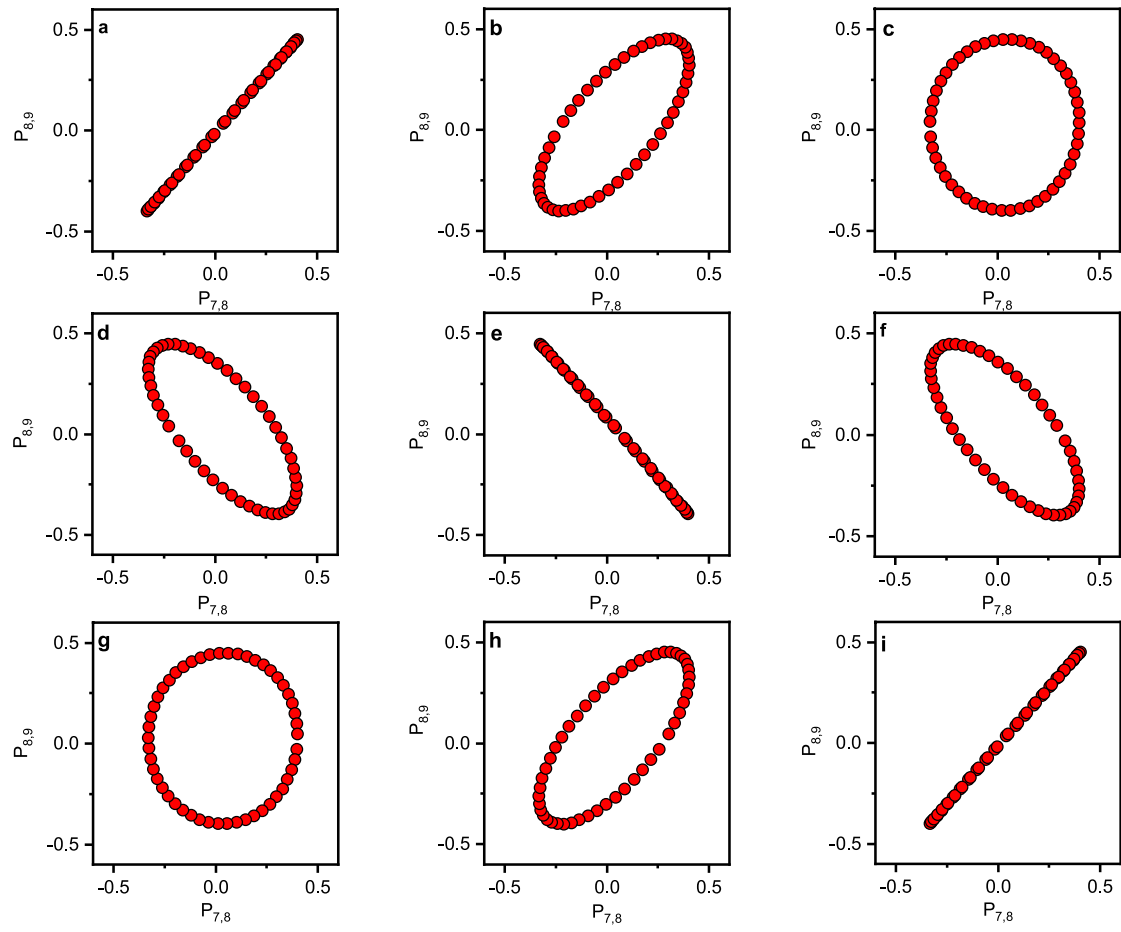
Peer review information *Nature* thanks Eduard Prat Costa, Makina Yabashi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

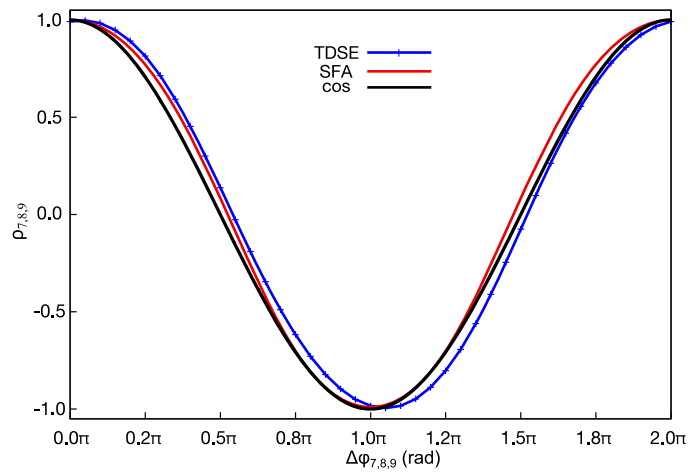


Extended Data Fig. 1 | Free-electron laser configuration for the generation of multiple harmonics and experimental setup. a, b, Configurations of the six undulators (U_1 – U_6) for the generation of three (**a**) and four (**b**) harmonics. In the first case, two undulators per harmonic were used, while in the second case, each harmonic was generated by one undulator only. The phase shifters (PS_1 – PS_5) used to control the relative phase between the harmonics are indicated in yellow for the two configurations. **c, d,** Typical single-shot photoelectron spectra without (black lines) and with (red lines) the NIR pulse,

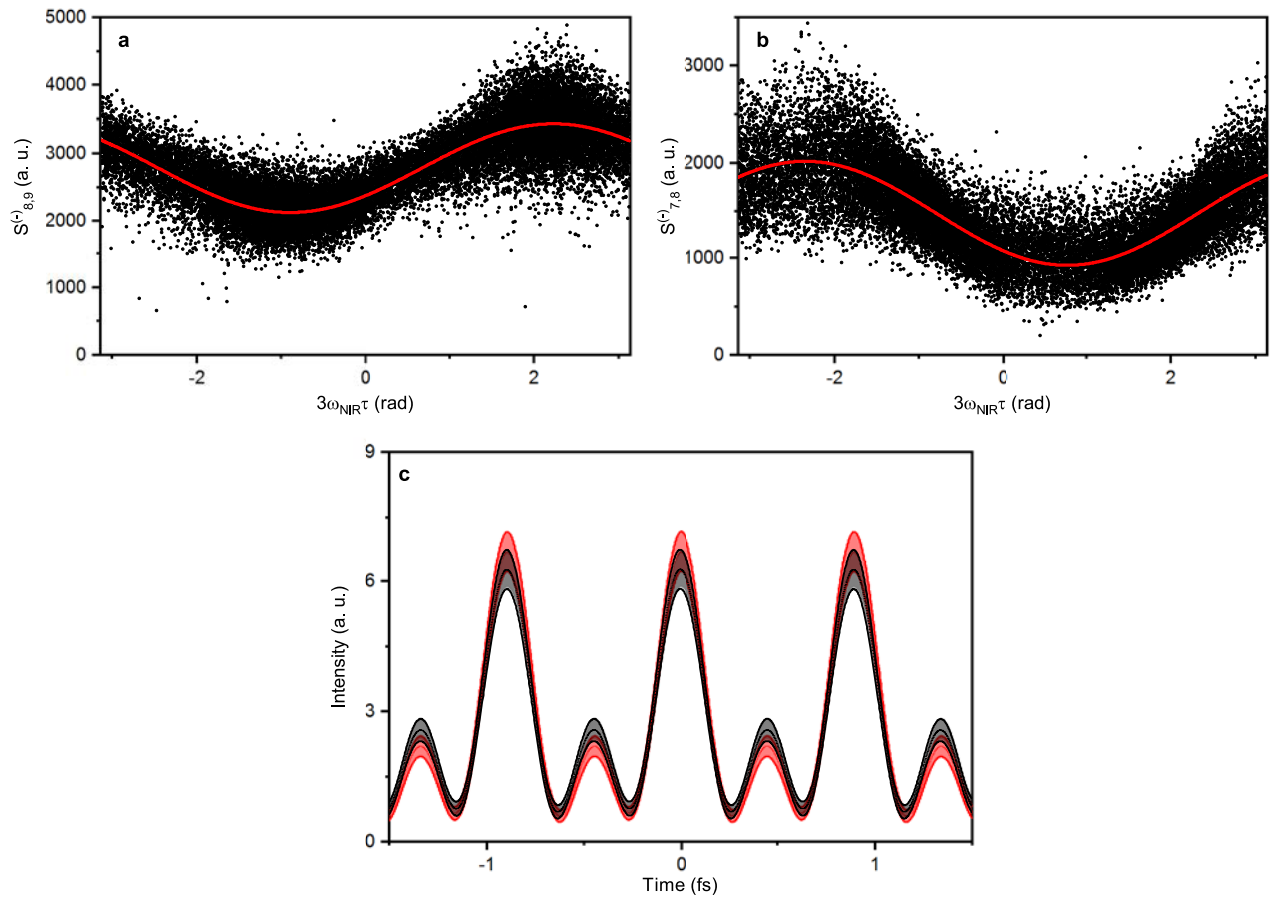
measured for the three (**c**) and four (**d**) harmonic cases. **e,** Schematic, half-section view of the spectrometer, including the ion flight tube (bottom) and electron flight tube (top). **f,** Normalized simulated geometrical collection efficiency as a function of polar emission angle for 2–42 eV electrons, using a cylindrical magnet configuration with the pole placed 5 mm away from the interaction region. Electrons were simulated using steps of 5 eV. An emission angle of 0° (180°) corresponds to the axis of the spectrometer in (away from) the direction of the electron detector.



Extended Data Fig. 2 | Simulated correlation plots. Shown are simulated correlation plots ($P_{8,9}$, $P_{7,8}$) for different values of $\Delta\phi_{7,8,9}$ from 0 to 2π in steps of $\pi/4$: $\Delta\phi_{7,8,9} = 0$ (a), $\pi/4$ (b), $\pi/2$ (c), $3\pi/4$ (d), π (e), $5\pi/4$ (f), $3\pi/2$ (g), $7\pi/4$ (h) and 2π (i). The intensities of the three harmonics are equal.

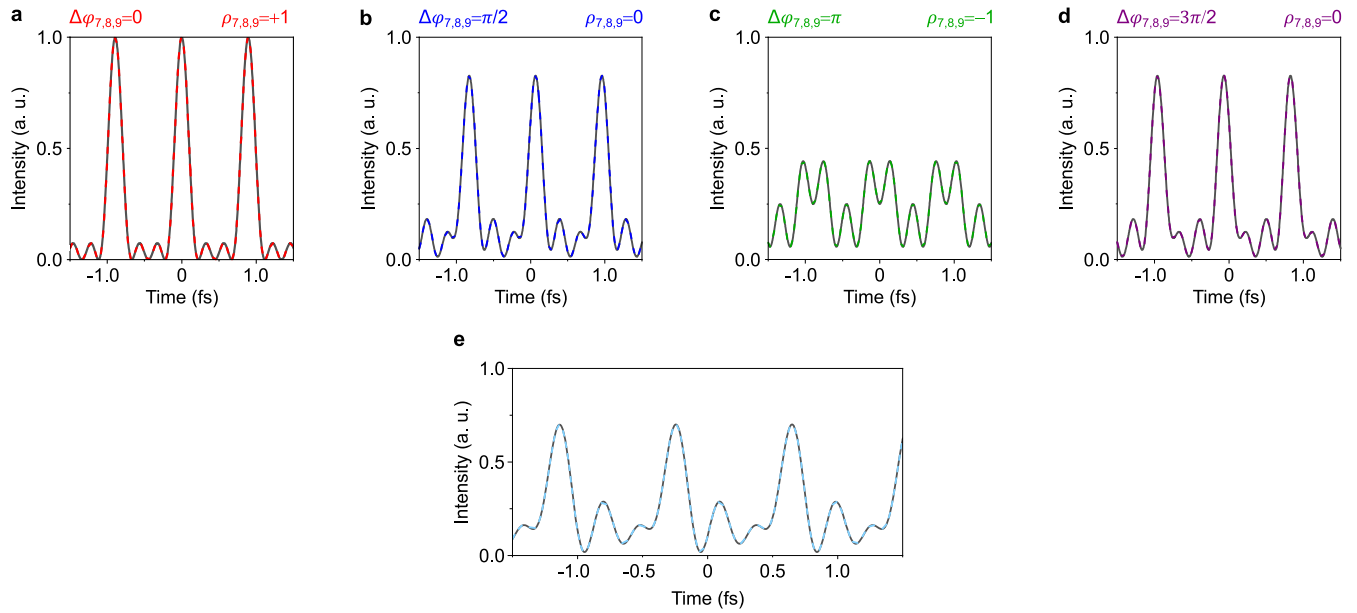


Extended Data Fig. 3 | Simulated correlation parameter $\rho_{7,8,9}$. Evolution of the correlation parameter $\rho_{7,8,9}$ as a function of the phase difference $\Delta\varphi_{7,8,9}$ simulated using the SFA (red) and the TDSE (blue). The black curve indicates a cosine evolution.



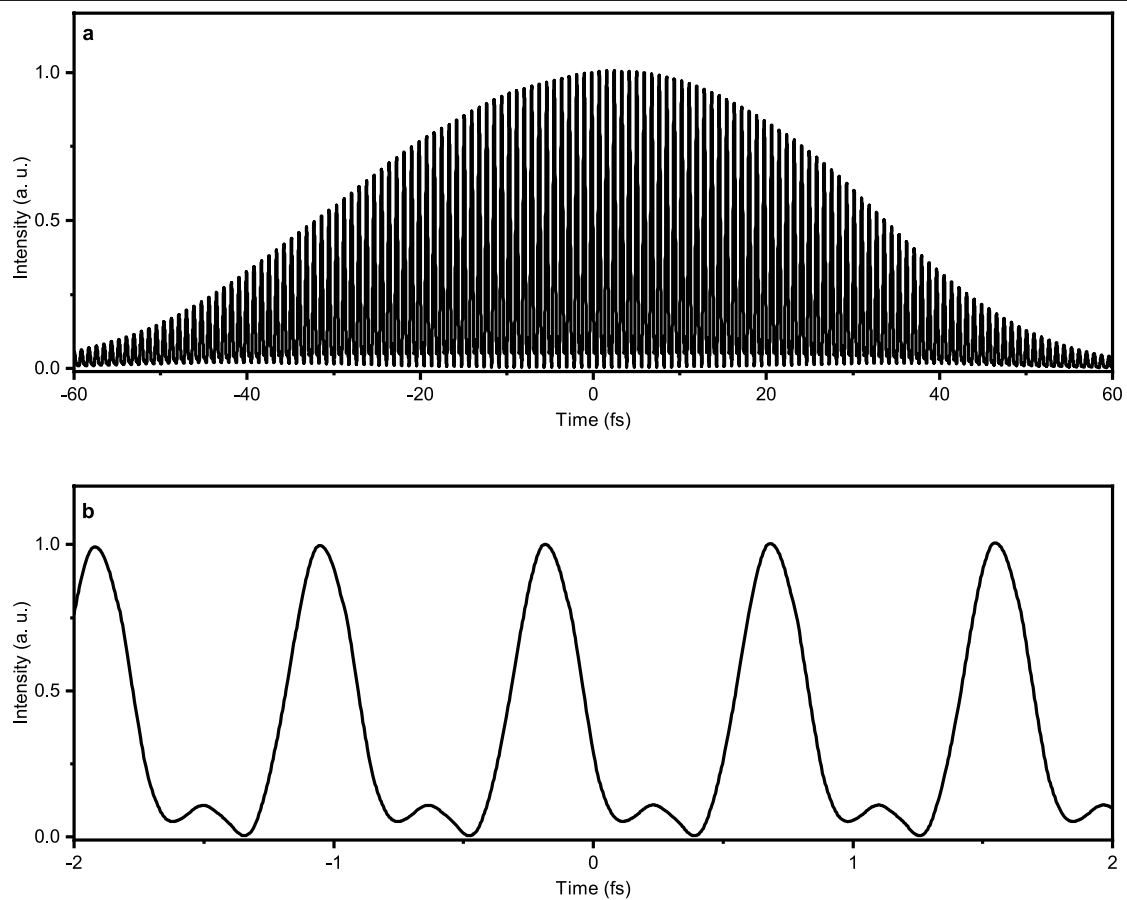
Extended Data Fig. 4 | Phase reordering of single-shot sideband intensities. **a, b,** Intensity of the sidebands $S_{8,9}^{(-)}$ (**a**) and $S_{7,8}^{(-)}$ (**b**) (black points) as a function of the relative phase $3\omega_{\text{NIR}}\tau$ between the attosecond pulse train and the NIR field. The red curves show sinusoidal fits of the distributions. **c,** Comparison of the reconstructed attosecond pulse train using the correlation parameter method

$p_{7,8,9}$ (black curve) and the 'reconstruction of attosecond beating by interference of two-photon transitions' method (red curve) based on the phase differences extracted from the sinusoidal fits. The second method is typically used for the characterization of attosecond pulse trains produced by HHG. The error in the reconstructions is indicated by the shaded areas.



Extended Data Fig. 5 | Reconstruction of attosecond pulses for multi-NIR photon transitions. **a–d**, Input (black line) and reconstructed (**a**, red line; **b**, blue line; **c**, green line; and **d**, magenta line) intensity profiles of the attosecond train, corresponding to Fig. 1c–f for phase differences $\Delta\varphi_{7,8,9}=0$ (**a**), $\pi/2$ (**b**), π (**c**) and $3\pi/2$ (**d**). **e**, Reconstruction of attosecond pulses from

sideband oscillations for multi-NIR photon transitions for the trace presented in Fig. 1b (input (black line) and reconstructed (blue dotted line) intensity profiles). The intensity of the NIR pulse is $I_{\text{NIR}} = 1.5 \times 10^{11} \text{ W cm}^{-2}$. The relative phases between the harmonics are: $\varphi_{10} - \varphi_9 = 108^\circ$, $\varphi_9 - \varphi_8 = 160^\circ$ and $\varphi_8 - \varphi_7 = 8^\circ$.



Extended Data Fig. 6 | GENESIS1.3 simulations. Shown is the attosecond pulse train simulated using the GENESIS1.3 code: **a**, complete temporal evolution of the train, and **b**, magnified view of the attosecond pulses in the train.

Extended Data Table 1 | XUV experimental parameters

	Harmonic	Photon energy (eV)	Energy (μJ)	Duration (fs)	Intensity (W/cm ²)
Three harmonics	7	32.88 ± 0.03	4.21 ± 0.58	50 ± 5	(1.1 ± 0.2) × 10 ¹⁴
	8	37.57 ± 0.04	5.29 ± 0.58	50 ± 5	(1.5 ± 0.2) × 10 ¹⁴
	9	42.27 ± 0.04	6.7 ± 1.2	50 ± 5	(1.9 ± 0.4) × 10 ¹⁴
Four harmonics	7	32.88 ± 0.03	1.01 ± 0.12	50 ± 5	(2.8 ± 0.4) × 10 ¹³
	8	37.57 ± 0.04	0.95 ± 0.12	50 ± 5	(2.6 ± 0.4) × 10 ¹³
	9	42.27 ± 0.04	0.67 ± 0.11	50 ± 5	(1.9 ± 0.4) × 10 ¹³
	10	46.96 ± 0.05	0.40 ± 0.50	50 ± 5	(1.1 ± 0.3) × 10 ¹³

Measured harmonic order, photon energy, energy of the single harmonic and intensity of the single harmonic for the three- and four-harmonic experiments. For the duration of the single harmonic, we used the values reported in ref. ³².

Extended Data Table 2 | Experimental correlation coefficients

Panel	a	b	c	d	e
$\rho_{7,8,9}$	0.83 ± 0.02	0.74 ± 0.01	0.32 ± 0.03	-0.20 ± 0.02	-0.62 ± 0.03
$\Delta\phi_{7,8,9}$ (rad)	-0.21 ± 0.12	0.50 ± 0.05	1.21 ± 0.05	1.91 ± 0.03	2.62 ± 0.05

Panel	f	g	h	i	j
$\rho_{7,8,9}$	-0.75 ± 0.02	-0.46 ± 0.01	0.14 ± 0.03	0.67 ± 0.01	0.86 ± 0.01
$\Delta\phi_{7,8,9}$ (rad)	3.33 ± 0.20	4.04 ± 0.03	4.74 ± 0.03	5.45 ± 0.02	6.16 ± 0.02

Correlation coefficients $\rho_{7,8,9}$ and phase differences $\Delta\phi_{7,8,9}$ for the measurements presented in the panels of Fig. 2 in the main text.

Extended Data Table 3 | Amplitudes and harmonic phase differences

	Figure	F_{10}	F_9	F_8	F_7	$\Delta\varphi_{8,9,10}$ (rad)	$\Delta\varphi_{7,8,9}$ (rad)
Three harmonics	3b,e	-	0.95 ± 0.06	1.00 ± 0.03	0.89 ± 0.06	-	0.08 ± 0.08
	3c,f	-	0.94 ± 0.07	1.00 ± 0.03	0.87 ± 0.06	-	1.93 ± 0.03
	3d,g	-	0.92 ± 0.07	1.00 ± 0.03	0.88 ± 0.06	-	3.29 ± 0.24
	3i,l	-	0.75 ± 0.06	1.00 ± 0.04	1.06 ± 0.06	-	6.08 ± 0.07
	3j,m	-	0.78 ± 0.08	1.02 ± 0.04	0.80 ± 0.06	-	6.09 ± 0.40
	3k,n	-	0.85 ± 0.09	0.83 ± 0.05	0.82 ± 0.07	-	6.18 ± 0.12
Four harmonics	4a-c	0.59 ± 0.07	0.76 ± 0.06	1.00 ± 0.05	1.03 ± 0.05	0.20 ± 0.15	5.94 ± 0.04
	4d-f	0.58 ± 0.07	0.77 ± 0.07	1.04 ± 0.05	1.07 ± 0.05	0.20 ± 0.15	1.23 ± 0.06
	4g-i	0.57 ± 0.08	0.77 ± 0.08	1.03 ± 0.06	0.99 ± 0.05	2.89 ± 0.08	2.95 ± 0.09

Amplitudes F_{10} , F_9 , F_8 and F_7 , and phase differences $\Delta\varphi_{8,9,10}$ and $\Delta\varphi_{7,8,9}$ for the three- (Fig. 3) and four-harmonic cases (Fig. 4). For the phase (amplitude) shaping in the three-harmonic case, the photoelectron spectra were rescaled to the area of the peak corresponding to the eighth harmonic in Fig. 3b and e (Fig. 3i and l). For the four-harmonic case, the photoelectron spectra were rescaled to the area of the eighth harmonic in Fig. 4a–c. For the pulse reconstruction, the measured $\Delta\varphi_{7,8,9}$ phases were corrected for the shift of 0.157 rad obtained by the TDSE simulations.

Extended Data Table 4 | Parameters for the GENESIS 1.3 simulations

Undulator parameters	
U ₁ -U ₂ resonant wavelength	28.889 nm
U ₃ -U ₄ resonant wavelength	32.500 nm
U ₅ -U ₆ resonant wavelength	37.143 nm
Undulator polarisation	linear
PS ₁ -PS ₂	0 rad
Dispersion	60 μm
Seed laser parameters	
Wavelength	260 nm
Power	40 MW
Pulse length (FWHM)	110 fs
Electron beam parameters	
Energy	1.2 GeV
Energy spread	110 keV
Normalized emittance	1 mm rad
Peak current	700 A
Beam size	50 - 70 μm

Parameters used in the GENESIS 1.3 code for simulating the generation of the train of attosecond pulses.

A droplet-based electricity generator with high instantaneous power density

<https://doi.org/10.1038/s41586-020-1985-6>

Received: 27 May 2019

Accepted: 6 November 2019

Published online: 5 February 2020

Wanghuai Xu^{1,2,9}, Huanxi Zheng^{1,9}, Yuan Liu^{3,4,9}, Xiaofeng Zhou^{1,9}, Chao Zhang¹, Yuxin Song¹, Xu Deng⁵, Michael Leung⁶, Zhengbao Yang¹, Ronald X. Xu², Zhong Lin Wang^{7*}, Xiao Cheng Zeng^{3,4*} & Zuankai Wang^{1,8*}

Extensive efforts have been made to harvest energy from water in the form of raindrops^{1–6}, river and ocean waves^{7,8}, tides⁹ and others^{10–17}. However, achieving a high density of electrical power generation is challenging. Traditional hydraulic power generation mainly uses electromagnetic generators that are heavy, bulky, and become inefficient with low water supply. An alternative, the water-droplet/solid-based triboelectric nanogenerator, has so far generated peak power densities of less than one watt per square metre, owing to the limitations imposed by interfacial effects—as seen in characterizations of the charge generation and transfer that occur at solid–liquid^{1–4} or liquid–liquid^{5,18} interfaces. Here we develop a device to harvest energy from impinging water droplets by using an architecture that comprises a polytetrafluoroethylene film on an indium tin oxide substrate plus an aluminium electrode. We show that spreading of an impinged water droplet on the device bridges the originally disconnected components into a closed-loop electrical system, transforming the conventional interfacial effect into a bulk effect, and so enhancing the instantaneous power density by several orders of magnitude over equivalent devices that are limited by interfacial effects.

Our droplet-based electricity generator (DEG) is based on our recent work¹⁹ showing that the continuous impinging of water droplets on a fluorinated material induces a high charge density on its surface. Our DEG device (Fig. 1a) is fabricated using drop-casting of polytetrafluoroethylene (PTFE), deposited with a tiny piece of aluminium, onto a glass substrate coated with indium tin oxide (ITO). As shown in Fig. 1b and Extended Data Fig. 1, the as-fabricated device is optically transparent, smooth and slippery. We hypothesized that, with continuous droplet impinging, the PTFE—a promising electret material with high charge-storage capability and stability^{20,21}—could serve as an ideal reservoir for charge storage, while electrostatically inducing opposite charge of the same amount on the ITO for possible charge transfer to an aluminium electrode. We find that when a falling water droplet spreads on the PTFE surface, it bridges the originally disconnected components (the PTFE/ITO and aluminium electrode) into a closed-loop, electrical system.

Figure 1c shows the time-dependent variation of measured surface charges on the PTFE film of an as-fabricated device under a relative humidity of 65.0%. With an increase in the number of impinging tap-water droplets (ion concentration 3.1 mM), there is a gradual increase in the amount of surface charge^{19,22}. Eventually, after around 1.6×10^4 droplets, the surface charge reaches a stable value of about 49.8 nC, indicating that continuous droplet impinging can serve as a robust way

to maintain stable and sufficient surface charge on the PTFE surface (Extended Data Fig. 2a).

We measured the electricity generation of an individual impinging droplet on the as-fabricated device in which the PTFE surface had been stored with sufficient and stable charges as a result of contact electrification between liquid and solid after continuous droplet impinging up to around 1.6×10^4 times. As shown in Fig. 1d and Supplementary Video 1, 400 commercial light-emitting diodes (LEDs) could be powered to instantaneously light up when four droplets of 100.0 μ l each, released from a height of 15.0 cm, contact the device. Focusing on an individual DEG indicates that the open-circuit output voltage and short-circuit current were about 143.5 V (Fig. 1e) and 270.0 μ A (Fig. 1f), respectively—around 295.0 and 2,600.0 times higher than the values obtained without an aluminium electrode (Extended Data Figs. 2b, 3). The instantaneous peak power density is 50.1 W m^{−2} under a load resistance of 332.0 k Ω (Extended Data Fig. 2c), which is three orders of magnitude higher than that of the control device without an aluminium electrode. We calculate the average energy-conversion efficiency of our DEG—defined as the harvested electrical energy relative to the input energy of an impinging droplet—to be roughly 2.2%, which is several orders of magnitude higher than that of our control device without an aluminium electrode. Note that the instantaneous peak density can be enhanced further by increasing the surface

¹Department of Mechanical Engineering, City University of Hong Kong, Hong Kong, China. ²Department of Precision Machinery and Precision Instrumentation, University of Science and Technology of China, Hefei, Anhui, China. ³Department of Chemistry, University of Nebraska-Lincoln, Lincoln, NE, USA. ⁴Department of Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE, USA. ⁵Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, China. ⁶School of Energy and Environment, City University of Hong Kong, Hong Kong, China. ⁷Beijing Institute of Nanoenergy and Nanosystems, Chinese Academy of Sciences, Beijing, China. ⁸Shenzhen Research Institute of City University of Hong Kong, Shenzhen, China. ⁹These authors contributed equally: Wanghuai Xu, Huanxi Zheng, Yuan Liu, Xiaofeng Zhou. *e-mail: zhong.wang@mse.gatech.edu; xzeng1@unl.edu; zuankaiwang@cityu.edu.hk

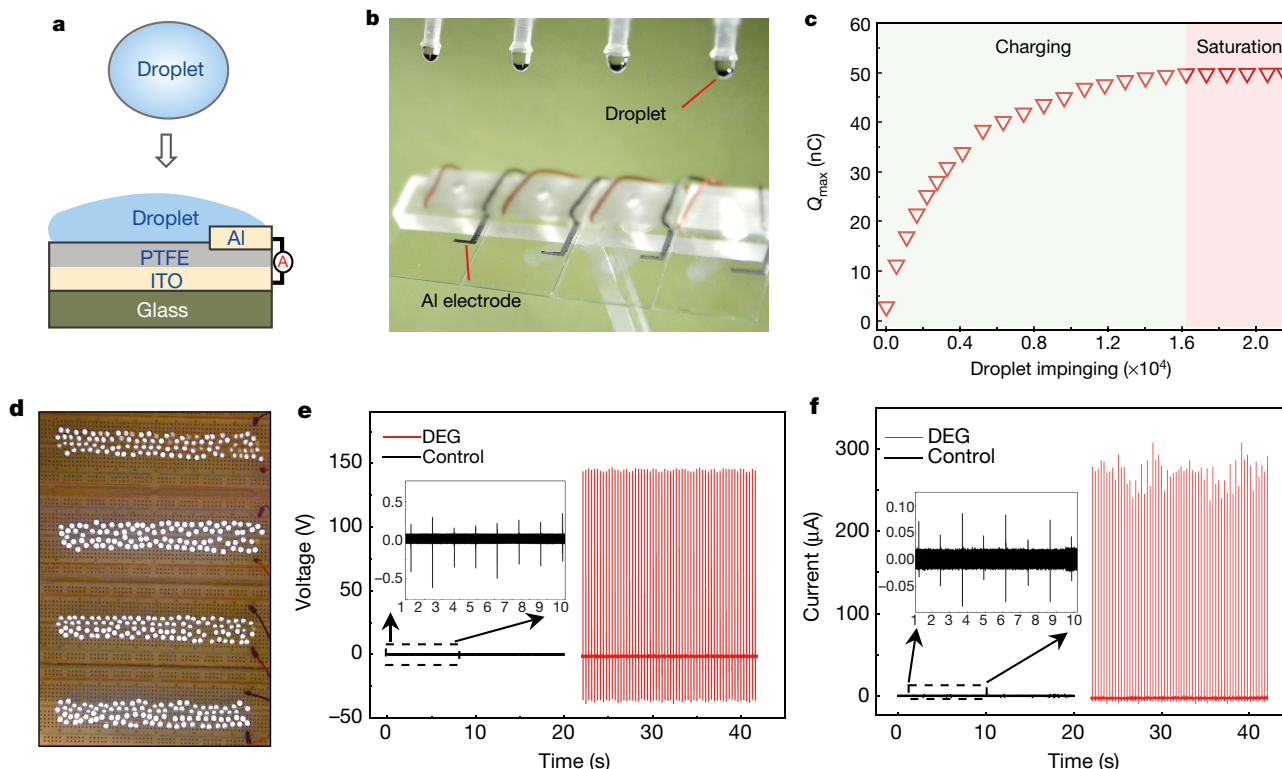


Fig. 1 | Design of the DEG. **a**, Schematic diagram. **b**, Optical image showing four parallel DEG devices fabricated on a glass substrate. The volume of each droplet is 100.0 μl . **c**, As individual droplets continue to impinge on the as-fabricated device, the amount of charge on the PTFE surface increases gradually and eventually reaches a stable value. **d**, One hundred commercial LEDs can be powered when one droplet, released from a height of 15.0 cm, is in contact with the device. **e**, Under the same experimental conditions (for example, the same droplet size and height of release), the output voltage measured from the DEG (in red, with the frequency of impinging droplets being

set at 4.2 Hz, and the total number of droplets being about 84) is more than two orders of magnitude higher than that from the control device (in black, with a droplet frequency of 1.0 Hz, and a total of 20 impinging droplets). The negligible electricity generation from the control device is limited by the interfacial effect, although its PTFE surface is loaded with the same amount of charge as the DEG. **f**, Comparison of output current from the DEG (in red) and the control device (in black) in response to continuous impinging of individual droplets.

charge of the PTFE film using an ion-injection method^{23,24} (Extended Data Fig. 4a, b). However, the long-term operation of the DEG device precharged by ion injection is susceptible to a gradual degradation of surface charge, eventually exhibiting performance comparable to that obtained through continuous droplet impact (Extended Data Fig. 4c).

The boost in the output performance of our device compared with the conventional design suggests that the DEG might operate via a different mechanism. First, as shown in Fig. 2a, the essential charges carried by the droplet before and after its impact on the precharged DEG are negligible. Moreover, there is no notable difference in the charge generated by droplets dispensing from a grounded versus an ungrounded outlet²⁵. Having ruled out any effects of droplets themselves and of the dropper on electricity generation, we next analysed the time-dependent evolution of the output current (Fig. 2b, c). Initially, upon contact with the PTFE surface, there is no apparent output current from the spreading droplet: the device is in a ‘switched-off’ state. This is essentially what occurs in the conventional design, with the charge generation being limited by an interfacial effect. The current then exhibits a large acceleration with a pronounced peak of up to 213.7 μA at an on-time (t_{on}) of 0 ms, transitioning into a switched-on state. Careful inspection shows that the sharp increase in the current originates from the contact of the spreading droplet with the aluminium electrode. We propose that this is a result of directional and rapid transfer of charge from the ITO electrode to the aluminium electrode. As plotted in Fig. 2d, in the early stage of droplet spreading, there is a rapid increase in the measured charges—a pattern consistent with the observed current–time curve. As the droplet continues to spread, charge transfer between

the ITO and aluminium electrodes continues until the droplet reaches its maximum spreading area, A_{max} , of 2.7 cm^2 , which is associated with a maximum charge, Q_{max} , of 49.8 nC (Fig. 2d). With retraction and sliding of the droplet from the impacting centre, the positive current turns to negative, indicating a back flow of charge from the aluminium electrode to the ITO. At an off-time (t_{off}) of 16.0 ms, the water droplet can be fully detached from the slippery aluminium electrode; this is accompanied by the output current and charge dropping to zero (Supplementary Video 2). In this condition, all charge is restored to the ITO, and a new cycle starts. This reversibility is confirmed by the measurement of cyclic charge. As shown in Fig. 2d, the amount of charge transferred between the ITO and aluminium electrodes in each cycle is constant—an indication that there is no deterioration of surface charge on the PTFE film. This is also suggested by the long-term measurement of charge stability (Fig. 1c).

To further understand the mechanisms underlying the performance of our DEG, we next examined the variation in the measured Q_{max} transferred from the ITO to the aluminium electrode as a function of the Weber number (Extended Data Fig. 5). This number is defined as $We = \rho D v^2 / \gamma$, where D , v and γ are respectively the diameter, impact velocity and surface tension of the droplet^{26–28}. As shown in Fig. 2e, with an increase in the Weber number from 7.7 to 150.4, the transferred charge increases from 13.3 nC to 53.3 nC. Thus, the increase in the amount of transferred charge between the ITO and aluminium electrodes in response to a varying Weber number suggests that the electricity generation is exquisitely regulated by the interaction between the impinging droplet and the configuration of the DEG, rather than

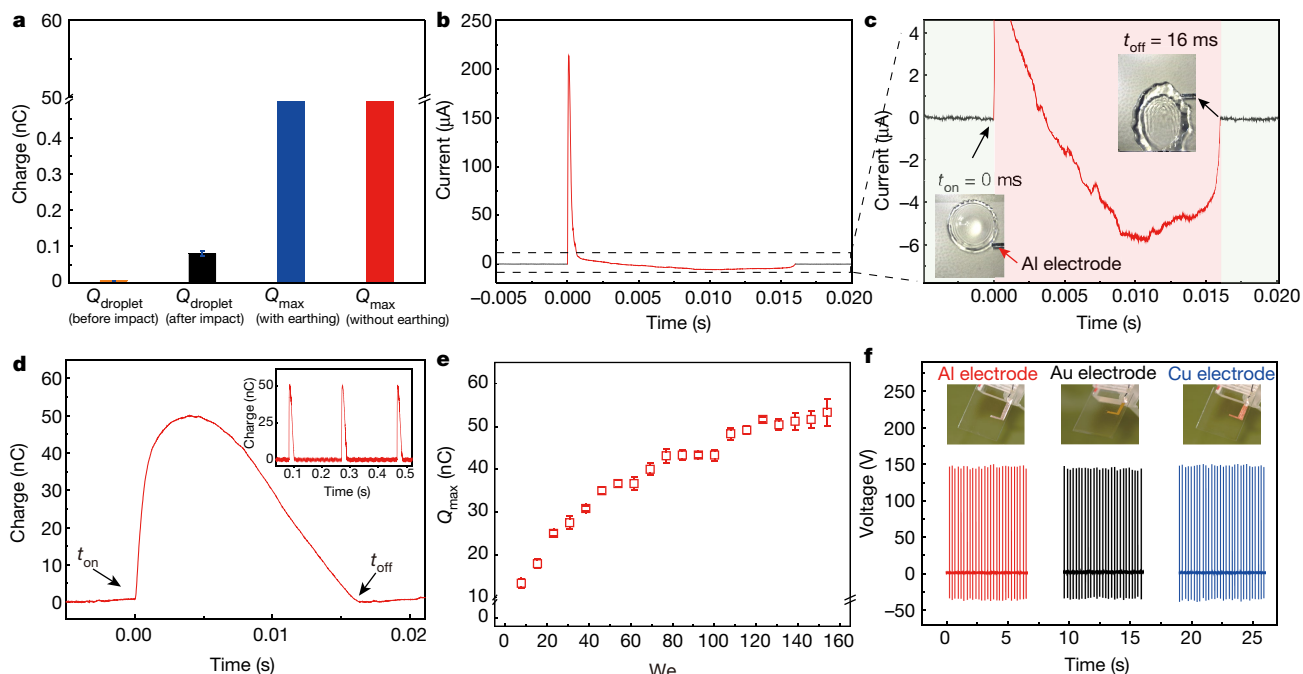


Fig. 2 | Origin of boosted electricity generation. **a**, The charges carried by the droplet (Q_{droplet}) before and after impinging on the PTFE surface are negligible compared with the measured charge of the DEG (Q_{max}). Data are means \pm s.e.m. For each mean, the total number of measurements is around ten. **b**, Time-resolved variation in current generated from the DEG during the entire droplet impact process. The dashed lines delineate the specific part of the current waveform shown in **c**. **c**, Synchronization of droplet-spreading dynamics and current response, and mapping of the time-dependent variation in charge flowing between PTFE/ITO and the aluminium electrode. The droplet retracts but still maintains contact with the aluminium electrode, while the current reverses to a negative value.

Insets are snapshots showing droplet dynamics. **d**, Time-dependent variation in the transferred charge, Q , generated on the DEG by an impinging droplet, indicating that the charge can return to zero when the DEG moves to switched-off mode. **e**, Variation in the maximum charge, Q_{max} , generated by an impacting droplet on the DEG under different Weber numbers. Data are means \pm s.e.m. For each mean, the total number of measurements is around ten. **f**, Output voltages remain constant when the aluminium electrode is replaced by a gold or silver electrode, suggesting that electricity generation is not sensitive to the specific electrode material. For all specific electrodes, the frequencies of impinging droplets and the total numbers of droplets are 4.2 Hz and about 28, respectively.

originating from just interfacial contact electrification. Moreover, the output is insensitive to the size and spatial location of the aluminium electrode and to the electrode material (Fig. 2f and Extended Data Fig. 6).

Looking at the device from a circuit perspective, the spreading droplet can be treated as a resistor and the PTFE as a capacitor, C_p , in which the water/PTFE serves as the top plate and PTFE/ITO as the bottom plate. In the switched-off mode, no capacitor is formed at the water/

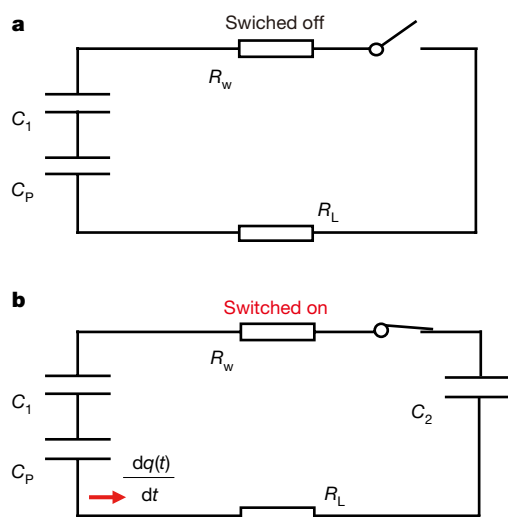
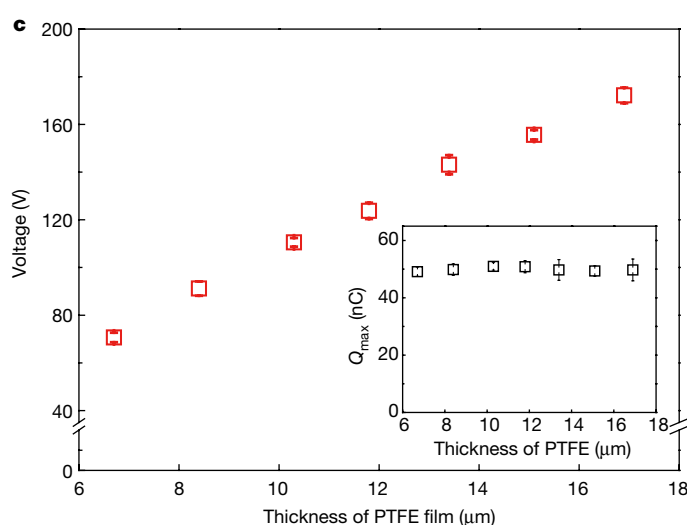


Fig. 3 | Circuit model. **a**, In the switched-off mode, there is no capacitor formed at the water/aluminium interface. As a result, C_p and C_1 remain in an open circuit and there is no charge flow between them. **b**, When the aluminium electrode and PTFE are connected by the water droplet (switched-on mode), another capacitor, C_2 , is established at the water/aluminium interface, forming a closed



circuit. R_w , R_L and $dq(t)/dt$ in the circuit are, respectively, the impedance of the water droplet, the impedance of the external load and the derivative of the transferred charge with respect to time. **c**, Output voltage and maximum charge (Q_{max}) as a function of PTFE thickness. Data are means \pm s.e.m. For each mean, the total number of measurements is ten.

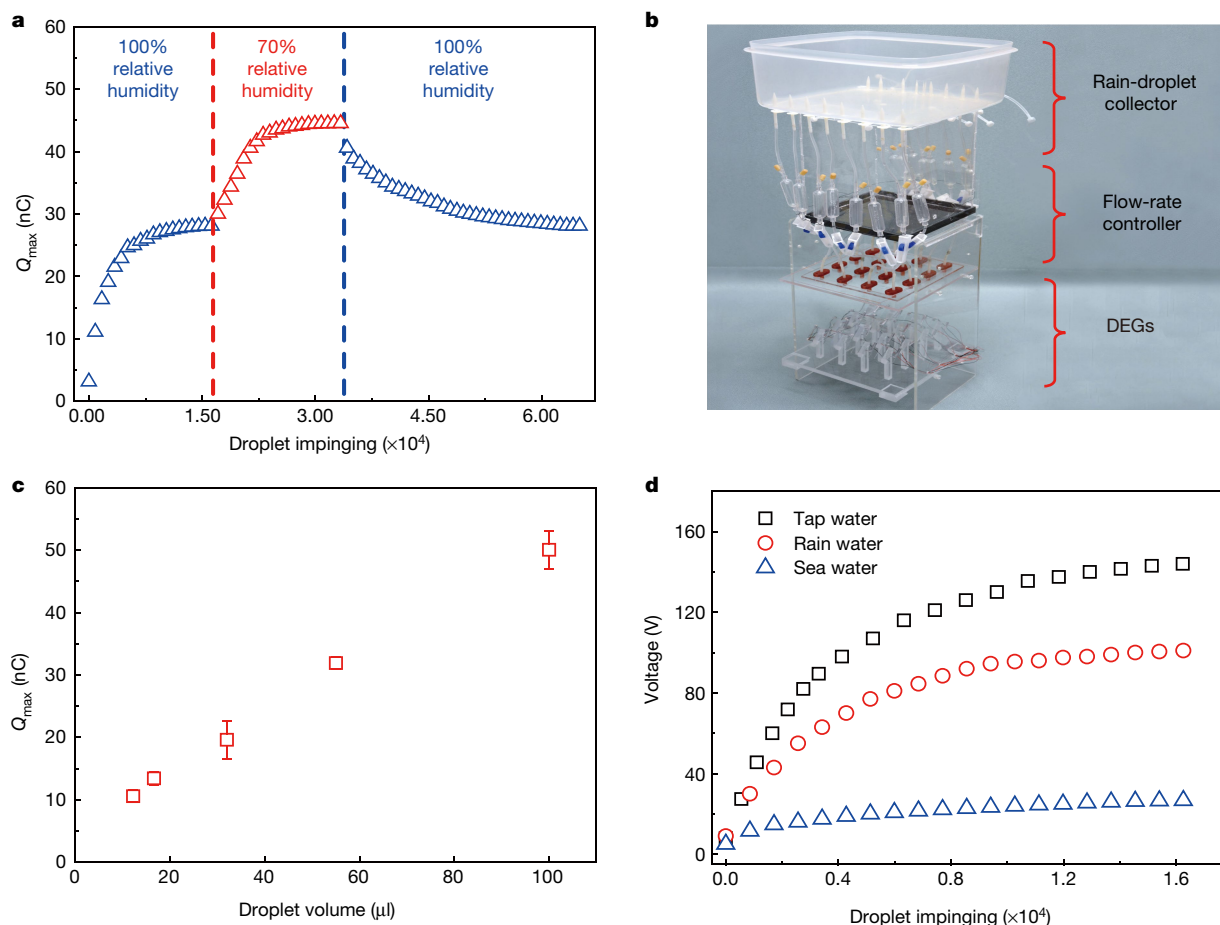


Fig. 4 | Stability and generality. **a**, Effect of relative humidity on charge loading stability. **b**, Optical image of our system for collecting and dispensing rainwater droplets. **c**, Tailoring the sizes of dispensing droplets for enhanced output.

Data are means \pm s.e.m. For each mean, the total number of measurements is around ten. **d**, Harvesting of hydrodynamic energy from different water sources: tap water, rainwater and sea water.

aluminium interface, and the circuit maintains an open state (Fig. 3a). By contrast, in the switched-on mode, capacitors are formed at the water/PTFE interface and the water/aluminium interface, transforming the original open circuit into a closed circuit (Fig. 3b). Given that the thickness of the PTFE is several orders of magnitude larger than that of the electric double layer at the water/solid interfaces, the capacitance of the capacitor C_p is negligible compared with that of the capacitors formed at the water/PTFE interface (C_1) and the water/aluminium interface (C_2). In combination with the high-density surface charge stored in the PTFE, the voltage across C_p is dramatically higher than that across C_1 and C_2 . Thus, the instantaneous peak output voltage, V , occurs when the spreading droplet is in contact with the aluminium electrode, and can be approximated as $Q_{\max}d/(\epsilon_p A_{\max})$, where d and ϵ_p are respectively the thickness and dielectric constant of the PTFE film. Given a measured A_{\max} of 2.7 cm² and a Q_{\max} of 49.8 nC, we calculate the voltage established across the PTFE to be roughly 143.5 V, consistent with our experimental measurement. Moreover, the measured peak voltage increases linearly with the thickness of the PTFE film, consistent with the predictions of our circuit model (Fig. 3c). Upon completion of the charging process, the C_p capacitor is recharged by the other two capacitors, as observed in current and charge measurements.

To gain molecular-level insights into the process of charge transfer between the PTFE and the aluminium electrode when a water droplet is in contact with both—that is, in the switched-on mode—we carried out molecular-dynamics simulations, using a nanoscale system with a water slab containing positive and negative ions. The simulations predict a

concerted motion and rapid separation of mobile positive and negative ions with the presence of an internal electric field between the PTFE and aluminium, once the fixed charges on the PTFE and aluminium are switched on (Extended Data Fig. 7). The separation of the mobile positive and negative charges in the water droplet towards the water/PTFE interface and the water/aluminium interface, respectively, reveals the charging process of the two capacitors at the water/PTFE interface (C_1) and water/aluminium interface (C_2) at the molecular level.

We also examined the stability of our devices under harsh environments involving high relative humidity²⁹. At a relative humidity of 100%, the surface charge increases gradually as the number of impinging droplets increases, reaching a saturated value of 28.1 nC (Fig. 4a). As the relative humidity is reduced to 70%, the surface charge increases rapidly and levels off at a stable value of 44.0 nC. On further exposure to 100% relative humidity, the charge output returns to 28.1 nC, indicating that continuous droplet impinging can help to maintain a constant and steady output even in harsh environments. This enhanced charge stability can be ascribed to the combination of continuous droplet impinging and the good charge-carrying capability of PTFE. By contrast, for control devices made of porous PTFE²¹, polydimethylsiloxane (PDMS) or polypropylene, under a relative humidity of 65.0%, performance decays owing to the poor charge stability of the surfaces³⁰ (Extended Data Fig. 8).

In addition to tap water, our DEG can harvest hydrodynamic energy from both raindrops and sea water. For raindrops, we designed a home-made platform consisting of a droplet collector and a

capillary-tubing-based dispenser (Fig. 4b and Supplementary Video 3). By adjusting the diameter of the capillary tubing and the height of release, we can precisely control the size and velocity of raindrops that contact the DEG arrays for enhanced on-demand output (Fig. 4c). Similarly, such a platform can separate a continuous flow of sea water into discontinuous droplet arrays, allowing for efficient electricity generation from a wide range of water-energy sources. We note that energy conversion from seawater droplets is lower than that from tap water and raindrops; however, it is still much higher than that of the conventional tap-water-based approach (Fig. 4d).

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-1985-6>.

- Lin, Z.-H., Cheng, G., Lee, S., Pradel, K. C. & Wang, Z. L. Harvesting water drop energy by a sequential contact-electrification and electrostatic-induction process. *Adv. Mater.* **26**, 4690–4696 (2014).
- Jeon, S.-B., Kim, D., Yoon, G.-W., Yoon, J.-B. & Choi, Y.-K. Self-cleaning hybrid energy harvester to generate power from raindrop and sunlight. *Nano Energy* **12**, 636–645 (2015).
- Xiong, J. et al. Wearable all-fabric-based triboelectric generator for water energy harvesting. *Adv. Energy Mater.* **7**, 1701243 (2017).
- Jin, S. et al. Large-area direct laser-shock imprinting of a 3D biomimic hierarchical metal surface for triboelectric nanogenerators. *Adv. Mater.* **30**, 1705840 (2018).
- Xu, W. H. et al. SLIPS-TENG: robust triboelectric nanogenerator with optical and charge transparency using a slippery interface. *Natl Sci. Rev.* **6**, 540–550 (2019).
- Yu, J., Ma, E. & Ma, T. Exponential energy harvesting through repetitive reconfigurations of a system of capacitors. *Commun. Phys.* **1**, 9 (2018).
- Wang, Z. L. New wave power. *Nature* **542**, 159–160 (2017).
- Zhu, G. et al. Harvesting water wave energy by asymmetric screening of electrostatic charges on a nanostructured hydrophobic thin-film Surface. *ACS Nano* **8**, 6031–6037 (2014).
- Scruggs, J. & Jacob, P. Harvesting ocean wave energy. *Science* **323**, 1176–1178 (2009).
- Dhiman, P. et al. Harvesting energy from water flow over graphene. *Nano Lett.* **11**, 3123–3127 (2011).
- Logan, B. E. & Elimelech, M. Membrane-based processes for sustainable power generation using water. *Nature* **488**, 313–319 (2012).
- Siria, A. et al. Giant osmotic energy conversion measured in a single transmembrane boron nitride nanotube. *Nature* **494**, 455–458 (2013).
- Feng, J. et al. Single-layer MoS₂ nanopores as nanopower generators. *Nature* **536**, 197–200 (2016).
- Schroeder, T. B. H. et al. An electric-eel-inspired soft power source from stacked hydrogels. *Nature* **552**, 214–218 (2017).
- Xue, G. et al. Water-evaporation-induced electricity with nanostructured carbon materials. *Nat. Nanotechnol.* **12**, 317–321 (2017).
- Huang, Y. et al. Interface-mediated hydroelectric generator with an output voltage approaching 1.5 volts. *Nat. Commun.* **9**, 4166 (2018).
- Chen, X. et al. Scaling up nanoscale water-driven energy conversion into evaporation-driven engines and generators. *Nat. Commun.* **6**, 7346 (2015).
- Nie, J. et al. Power generation from the interaction of a liquid droplet and a liquid membrane. *Nat. Commun.* **10**, 2264 (2019).
- Sun, Q. et al. Surface charge printing for programmed droplet transport. *Nat. Mater.* **18**, 936–941 (2019).
- Cui, L. et al. The comparative studies of charge storage stabilities among three PP/porous PTFE/PP electret. *J. Electrostat.* **67**, 412–416 (2009).
- Xia, Z., Gerhard-Multhaupt, R., Künstler, W., Wedel, A. & Danz, R. High surface-charge stability of porous polytetrafluoroethylene electret films at room and elevated temperatures. *J. Phys. D* **32**, L83 (1999).
- Yatsuzuka, K., Mizuno, Y. & Asano, K. Electrification phenomena of pure water droplets dripping and sliding on a polymer surface. *J. Electrostat.* **32**, 157–171 (1994).
- Zhao, P. et al. Emulsion electrospinning of polytetrafluoroethylene (PTFE) nanofibrous membranes for high-performance triboelectric nanogenerators. *ACS Appl. Mater. Interfaces* **10**, 5880–5891 (2018).
- Wang, S. et al. Maximum surface charge density for triboelectric nanogenerators achieved by ionized-air injection: methodology and theoretical understanding. *Adv. Mater.* **26**, 6720–6728 (2014).
- Thomson, W. On a self-acting apparatus for multiplying and maintaining electric charges, with applications to illustrate the voltaic theory. *Proc. R. Soc. Lond.* **16**, 67–72 (1868).
- Liu, Y. et al. Pancake bouncing on superhydrophobic surfaces. *Nat. Phys.* **10**, 515–519 (2014).
- Bird, J. C., Dhiman, R., Kwon, H.-M. & Varanasi, K. K. Reducing the contact time of a bouncing drop. *Nature* **503**, 385–388 (2013); erratum 505, 436 (2014).
- Richard, D., Clanet, C. & Quéré, D. Contact time of a bouncing drop. *Nature* **417**, 811 (2002).
- Nguyen, V. & Yang, R. Effect of humidity and pressure on the triboelectric nanogenerator. *Nano Energy* **2**, 604–608 (2013).
- Sessler, G. M. in *Electrets* (ed. Sessler, G. M.) 13–80 (Springer, 1980).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Methods

Materials

Acetone (RCI Labscan, 99.5%), ethanol (Sigma Aldrich, 97%), nitric acid (Sigma Aldrich, 70%), porous PTFE film (Sterlitech Corporation, PTU023001), PTFE precursor (Dupont AF 601S2, 6 wt%) and PDMS (Dow Corning Sylgard 184) were used without further purification. The PTFE precursor is composed of PTFE dissolved in 4,5-difluoro-2,2-bis(trifluoromethyl)-1,3-dioxole (a low-boiling organic solvent), which does not contain any extra additives.

Fabrication of DEG device

To fabricate the DEG, we first ultrasonically cleaned a piece of ITO glass slide, of size 30 mm × 30 mm × 0.4 mm, in acetone and then ethanol for 10 min each. We then deposited the PTFE precursor on the ITO glass by drop-casting, and heated it at 120 °C for 15 min to remove all solvent in the PTFE precursor. Upon curing at 120 °C, the PTFE precursor was transformed into a smooth and dense PTFE film, as shown by scanning electron microscopy (SEM; Extended Data Fig. 1). The thickness of the PTFE film can be adjusted by controlling the volume of precursor (Extended Data Fig. 9). To construct the aluminium electrode, we assembled a tiny conductive aluminium tape of size of 1 mm × 5 mm × 50 µm onto the as-prepared PTFE film. For comparison, we also fabricated a control device in the same way but without the aluminium electrode. To fabricate a control device with PDMS film as the dielectric layer, we spin-coated a liquid mix of polydimethylsiloxane and a curing agent (ratio 10/1) with a volume of 200 µl onto ITO glass at a speed of 3,000 revolutions per minute, and then cured the film at 80 °C for 1 h. To fabricate control devices with porous PTFE film and commercial polypropylene tape as the dielectric layer, we attached the porous PTFE film and commercial polypropylene tape directly onto the ITO glass slide.

Characterization and electrical measurement

We used a syringe pump and a plastic tube to generate water droplets. The droplet size could be tailored by varying the inner diameter of the plastic nozzle connecting to the outlet of the plastic tube. The inner diameter of the nozzle required to generate droplets of 100 µl was 6.0 mm. If not specified, the composition of water droplets was tap water at an ion concentration of 3.1 mM. The volume of water droplets was fixed at 100 µl and the droplet outlet was not earthed. We recorded the spreading and retraction dynamics of water droplets using a high-speed camera (Photron FASTCAMS4) at a typical recording speed of 6,000 frames per second. The voltage output of DEG was measured using an oscilloscope (Rohde and Schwarzrte, RTE1024) equipped with a high-impedance (10 MΩ) probe. We measured the current and the charges transferred between the ITO and aluminium electrode using the oscilloscope coupled with a low-noise current preamplifier (Stanford Research System Model SR570) and a Faraday cup connected with a nanocoulomb meter (Monroe model 284), respectively. The as-fabricated device was tilted at 45.0° for rapid liquid detachment. To measure the variation in maximum charges, Q_{\max} , transferred from ITO to aluminium as a function of the Weber number or the maximum spreading area, we varied the releasing heights of droplets between 1 cm and 20 cm. In typical measurements, we kept the relative humidity and the environmental temperature at approximately 65.0% and 20.0 °C, respectively.

Continuous droplet impinging and electricity generation

We showed in the main text that the DEG made of PTFE loaded with sufficient charge allows for reversible and efficient electricity generation. Here we demonstrate that sufficient charge on PTFE can be achieved by continuous droplet impinging. Extended Data Fig. 2a shows the variation in output voltage measured from an individual impinging droplet as a function of the number of droplets impinging. Q_{\max} and the output voltage increase gradually with increasing droplet impinging times,

eventually reaching a plateau with the charge and voltage stabilized at 49.8 nC and 143.5 V, respectively, after impinging of 1.6×10^4 times (Fig. 1c). This charge-loading method is applicable to a wide range of thicknesses of the PTFE film. Our result shows that the maximum transferred charges are comparable after 1.6×10^4 times of droplet impinging when the thickness of the PTFE film varies from 6.7 µm to 16.9 µm (Fig. 3c). Note that the voltage increases linearly with film thickness, agreeing with the predictions of our circuit model.

Comparison with a conventional generator

We also characterized the performance of a control sample that lacks the aluminium electrode. Note that the PTFE surface of this control was prepared using the same method as for DEG. Extended Data Fig. 3a, b shows an optical image of the as-fabricated droplet-based control device and a schematic drawing of its basic working mechanism^{1,4}. Before the droplet contacts the PTFE, the amount of (positive) charges on the ITO is the same as the (negative) charges on the PTFE because of electrostatic induction. Thus, there is no current flow from ITO to the ground (Extended Data Fig. 3b, i). When a water droplet leaves the PTFE surface after impacting (Extended Data Fig. 3b, ii), the droplet becomes positively charged while the PTFE is more negatively charged as a result of contact electrification (Extended Data Fig. 3b, iii)²². Accordingly, a flow of current between the ground and ITO electrode is induced (Extended Data Fig. 3b, iv). As shown in Extended Data Fig. 3c, d, the voltage and transferred charge generated from nine droplets impinging on the control device are measured when the frequency of impinging droplets is set at 1.0 Hz. For a single droplet, the output voltage and the amount of transferred charge are roughly −0.4 V and 0.075 nC, respectively, both of which are negligible compared with those of the DEG. Moreover, for the nine droplets, the total amount of transferred charge in the control device is measured to be identical with the accumulated charge carried by the departing droplet (Extended Data Fig. 3d), confirming that electricity generation in the control device indeed originates from the triboelectric effect, a natural interfacial phenomenon. Note that by using continuous droplet impinging or the ion-injection method^{23,24}, the amount of negative charge on the control PTFE surface can be enhanced, which can then induce positive charges on ITO. However, these positive charges cannot be released from the ITO because of attraction by negative charges on the PTFE, and there is no pronounced electricity generation, in striking contrast to the DEG. All of these results highlight the unique advantage of the DEG, which is characterized by a bulk effect and hence enhanced electricity generation.

Calculation of average conversion efficiency

To quantify the performance of our DEG, we calculated the average conversion efficiency of mechanical energy into electric energy. We first calculated the instantaneous conversion efficiency of the i th droplet impinging, η_i as follows:

$$\eta_i = \frac{E_{\text{out}}}{E_{\text{in}}} = \frac{\int_{t_{\text{on}}}^{t_{\text{off}}} \frac{U^2}{R_L} dt}{mgh} \quad (1)$$

where the droplet mass, m , is 0.1 g; the gravitational acceleration, g , is 9.8 m s^{−2}; the relative height between the releasing droplet and the DEG, h , is 15.0 cm; and $R_L = 10.0$ MΩ. For our DEG, the kinetic energy carried by a droplet (of 100 µl) released from a height of 15.0 cm at the 1.6×10^4 th droplet impinging (that is, $i = 1.6 \times 10^4$) is roughly 1.47×10^{-4} J, and the generated electrical energy is calculated as 3.2 µJ (Extended Data Fig. 4b), responding to a η_i of around 2.2%. Such efficiency is five orders of magnitude higher than that of our control device (around $2.1 \times 10^{-5}\%$).

Next, we discuss the overall conversion efficiency (η_A) of the first n droplets, including those droplets used in precharging; this conversion

efficiency is calculated as the average of the instantaneous conversion efficiency of all droplets ($i = 1, 2, 3, \dots, n$), expressed as:

$$\eta_A = \frac{\sum_{i=1}^n \eta_i}{n} \quad (2)$$

The η_i of individual droplets can be obtained as above (equation (1)). The overall efficiency will increase with the number of impacting droplets. For example, the overall efficiency approaches 2.2%—a conversion efficiency obtained for a single falling droplet impacting on the DEG in the steady state—when the number of impinging droplets is more than 2.0×10^5 .

Maximum spreading area

We also studied the dependence of the maximum spreading area, A_{\max} , on the surface charge under a fixed impacting Weber number of 100. For a PTFE film without loaded charges, A_{\max} is measured at 2.71 cm^2 , which is comparable with the A_{\max} value measured for PTFE with loaded charges (2.72 cm^2) (Extended Data Fig. 5). This result suggests that the maximum spreading area of a droplet on the PTFE surface relies on the predetermined release height of the impinging droplet, and is insensitive to the surface charge of the PTFE film.

Circuit analysis

We now discuss the entire droplet and device from a circuit perspective³¹. When a droplet spreads on a PTFE surface loaded with sufficient negative charge, the base contact area with the PTFE varies dynamically as a function of time. A capacitor, C_p , is formed with the water/PTFE as the top plate and PTFE/ITO as the bottom plate, respectively. At the water/PTFE interface, there is an additional capacitor, C_1 . Before the impinging droplet contacts the aluminium electrode, there is no capacitor formed at the water/aluminium interface. As a result, C_p and C_1 remain in an open circuit and there is no charge flow between them (Fig. 3a). By contrast, when the aluminium electrode and PTFE are connected by the liquid (switched-on mode), the other capacitor, C_2 , is established at the water/aluminium interface. Thus, C_1 , together with C_p and C_2 , form a close circuit. The instantaneous peak output voltage, V , occurs when the spreading droplet is in contact with the aluminium, and can be approximated as $Q_{\max}/(\epsilon_p A_{\max})$. In this circuit, the time-dependent capacitance of C_p , C_1 and C_2 can be expressed as $C_p(t) = A(t)\epsilon_p/d$, $C_1(t) = A(t)\epsilon_w/\lambda_{\text{EDL}}$ and $C_2(t) = A_1(t)\epsilon_w/\lambda_{\text{EDL}}$, respectively, where $A(t)$ and $A_1(t)$ are the time-dependent contact areas of the water/PTFE interface and the water/aluminium interface, respectively; and ϵ_w and λ_{EDL} are the dielectric constant of PTFE and the width of EDL, respectively. The equivalent circuit is shown in Fig. 3b, governed by the following differential equation:

$$(R_L + R_w) \frac{dq(t)}{dt} = \frac{Q_p(t) - q(t)}{C_p(t)} - \frac{q(t)}{C_1(t)} - \frac{q(t)}{C_2(t)}$$

$$q(t=0) = 0$$

where $q(t)$ is the transfer charge, and R_L and R_w are the impedance of the external load and water droplet, respectively.

Molecular-dynamics simulations

To simulate ion transport and separation in a water droplet in contact with the as-fabricated device, we carried out molecular-dynamics simulations. To this end, we used the transferable intermolecular potential with four points/for the simulation of water solid ice (TIP4P/ICE)³² water model, which is popular in molecular-dynamics simulations of water. Various properties of water—including static properties such as the melting point of ice, liquid density in ambient conditions, and water/ice phase diagram, as well kinetic properties such as water's diffusion

constant in ambient conditions—have been successfully reproduced using TIP4P/ICE. To mimic ion conduction in tap water, we introduce identical amounts of sodium (Na^+) and chloride (Cl^-) ions into the water. The molecular-dynamics system includes a slab of water containing 4.0×10^4 water molecules with 808 Na^+ and 808 Cl^- ions. To mimic the charged PTFE and aluminium electrode, we use rigid atomic trilayers. We fix 800 negative and 800 positive charges, with a spacing of 8.7 \AA , on the middle and bottom atomic layers of PTFE (for the negative charges) and the ITO electrode (for the positive charges); each site is charged $+e$ or $-e$. The box size of the model system is $17.3 \text{ nm} \times 17.3 \text{ nm} \times 31.4 \text{ nm}$, in which the thickness of water layer is about 4.5 nm . Periodic boundary conditions are applied in the x and y directions. The parameters for Na^+ and Cl^- are taken from previous work³³ ($\sigma_{\text{Na}} = 2.876 \text{ \AA}$; $\epsilon_{\text{Na}} = 0.5216 \text{ kJ mol}^{-1}$; $\sigma_{\text{Cl}} = 3.785 \text{ \AA}$; $\epsilon_{\text{Cl}} = 0.5216 \text{ kJ mol}^{-1}$; where ϵ is the depth of a potential well, and σ is the finite distance at which the interparticle potential is zero. The cross Lennard–Jones interaction parameters between water and Na^+ and Cl^- ions are given by the Lorentz–Berthelot rule. The interactions between substrate atoms (sub) and the NaCl water solution are described by the 12-6 Lennard–Jones potential ($\sigma_{\text{Na-sub}} = 3.021 \text{ \AA}$; $\epsilon_{\text{Na-sub}} = 0.4785 \text{ kJ mol}^{-1}$; $\sigma_{\text{Cl-sub}} = 3.476 \text{ \AA}$; $\epsilon_{\text{Cl-sub}} = 0.4785 \text{ kJ mol}^{-1}$; $\sigma_{\text{O-sub}} = 3.458 \text{ \AA}$; $\epsilon_{\text{O-sub}} = 0.6223 \text{ kJ mol}^{-1}$). We used the fast smooth particle-mesh Ewald method to model electrostatic interactions with a real-space cut-off of 10 \AA . Van der Waals interactions are truncated at 10 \AA . We integrate Newton's equations of motion with a time step of 1 fs by using a leap-frog algorithm in the molecular-dynamics simulations. We use a Nosé–Hoover scheme to maintain the systems at a constant temperature. All molecular-dynamics simulations are carried out using Gromacs 4.5.5 software. First, we perform molecular-dynamics simulations of water with dissolved Na^+ and Cl^- ions but without any charges on the substrates while the temperature is maintained at 300 K . The simulation lasts 3 ns . Next, we perform molecular-dynamics simulations in the constant-temperature and constant-volume ensemble at 300 K for both switched-off and switched-on mode. The simulation lasts 5 ns for each mode.

Data availability

The data that support the findings of this study are available from the corresponding authors on reasonable request.

31. Moon, J. K., Jeong, J., Lee, D. & Pak, H. K. Electrical power generation by mechanically modulating electrical double layers. *Nat. Commun.* **4**, 1487 (2013).
32. Abascal, J., Sanz, E., García Fernández, R. & Vega, C. A potential model for the study of ices and amorphous water: TIP4P/ice. *J. Chem. Phys.* **122**, 234511 (2005).
33. Koneshan, S. R., Rasaiah, J. C., Lynden-Bell, R. M. & Lee, S. H. Solvent structure, dynamics, and ion mobility in aqueous solutions at 25°C . *J. Phys. Chem. B* **102**, 4193–4204 (1998).

Acknowledgements Z.W. acknowledges financial support from the National Natural Science Foundation of China (grant 31771083), the Research Grants Council of Hong Kong (grants 11219219, C1018-17G and 11218417), the Shenzhen Science and Technology Innovation Council (grant JCYJ20170413141208098), the Innovation and Technology Council (grants 9440175), and the City University of Hong Kong (grants 9680212 and 9600011). X.C.Z. was supported by the United States National Science Foundation (CHE-1665324). X.Z. was supported by the Science and Technology Committee of the Shanghai Municipality (grant 1951120100) and by Initial Funding for Scientific Research of East China Normal University. Molecular-dynamics simulations were carried out using the computer facility at the University of Nebraska-Lincoln Holland Computing Center.

Author contributions Z.W. and W.X. conceived the research. W.X., H.Z., X.Z. and C.Z. designed the experiments. W.X. prepared the samples. W.X., H.Z. and Y.S. carried out the experiments. Y.L. and X.C.Z. conducted the molecular-dynamics simulations. All authors analysed the data. Z.W., R.X.X., X.D., X.C.Z., W.X., Z.L.W., X.Z., Y.L., M.L., Z.Y. and H.Z. wrote the manuscript.

Competing interests The authors declare no competing interests.

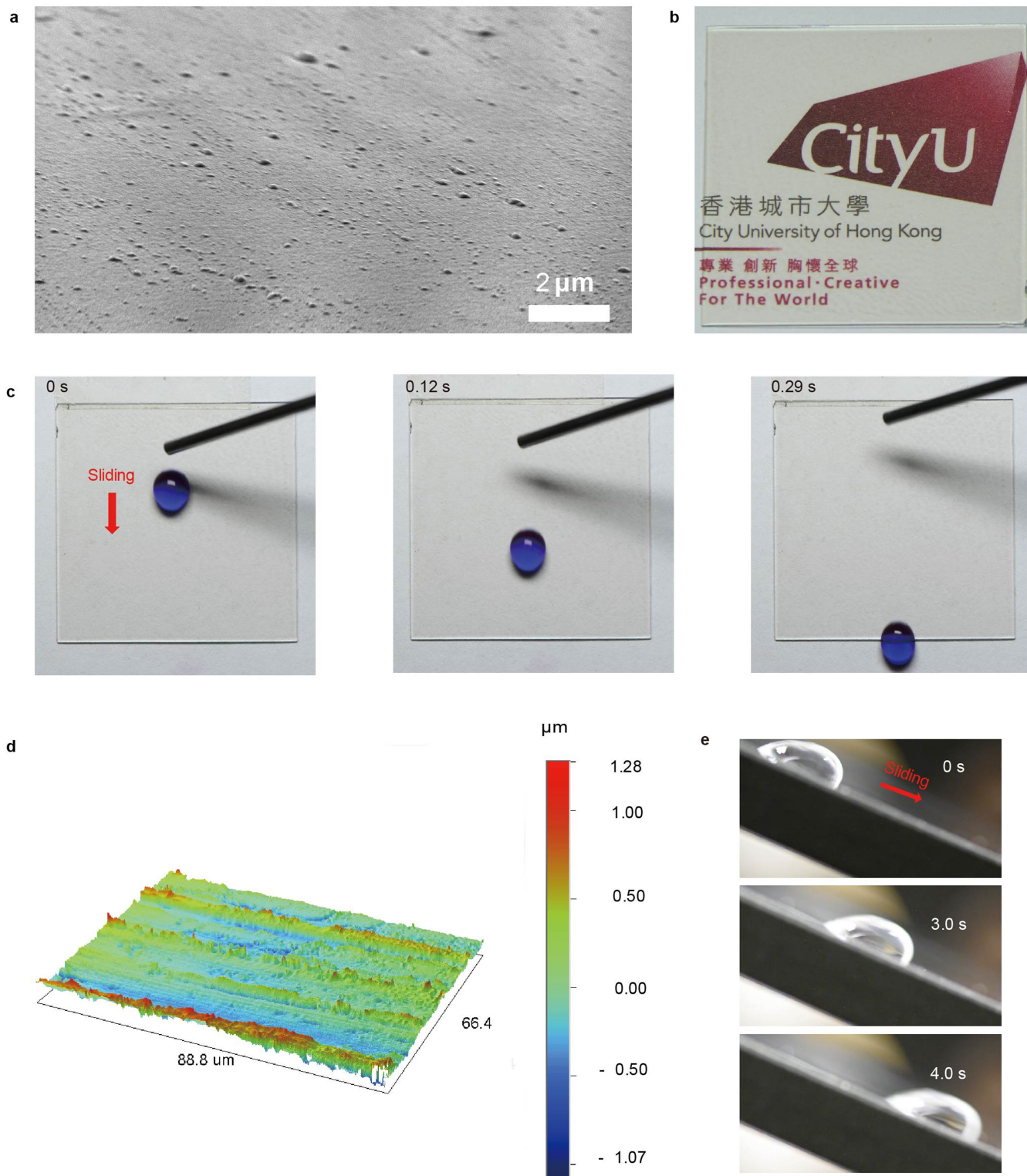
Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-1985-6>.

Correspondence and requests for materials should be addressed to Z.L.W., X.C.Z. or Z.W.

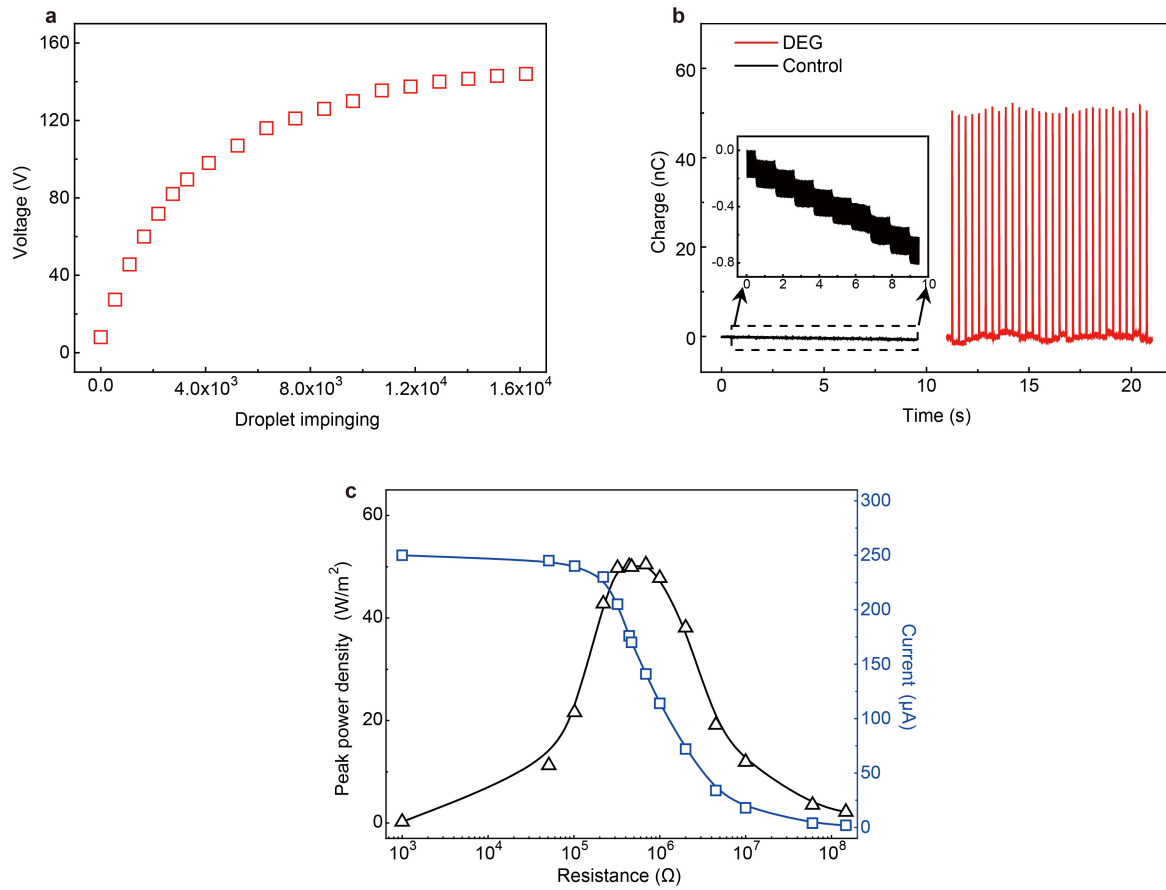
Peer review information Nature thanks Lars Egil Helseth and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



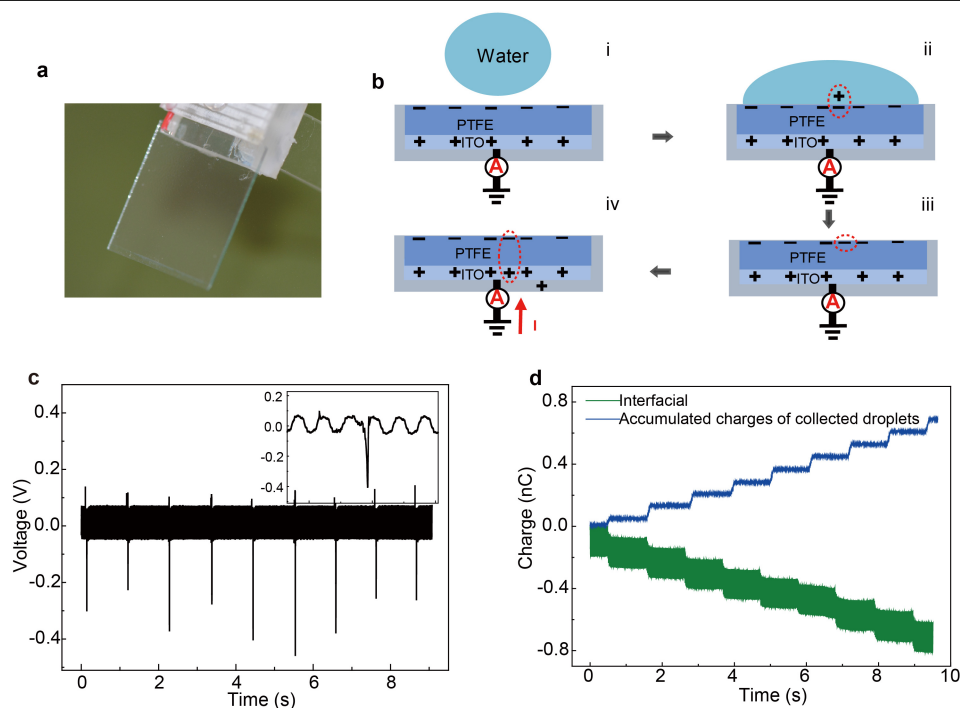
Extended Data Fig. 1 | Surface morphology and sliding behaviour of water droplets on the PTFE film and aluminium electrode. **a**, SEM image of the PTFE film used in our DEG. Upon curing at 120 °C and solvent evaporation, the PTFE precursor is transformed into a smooth and dense PTFE film. **b**, Photograph of the fabricated PTFE film together with the ITO glass on a logo, showing the high transparency of the film. **c**, A droplet of roughly 30.0 μl can easily slide on the

surface of a PTFE film made from pure PTFE solution (placed on a substrate with a tilt angle of 15.0°). **d**, Contour graph image of an aluminium electrode shows that its surface is very flat and uniform. **e**, A droplet can slide off an aluminium electrode without leaving residual water. The aluminium electrode is placed on a substrate with a tilt angle of 25.0°. There is no residual water on the electrode surface.



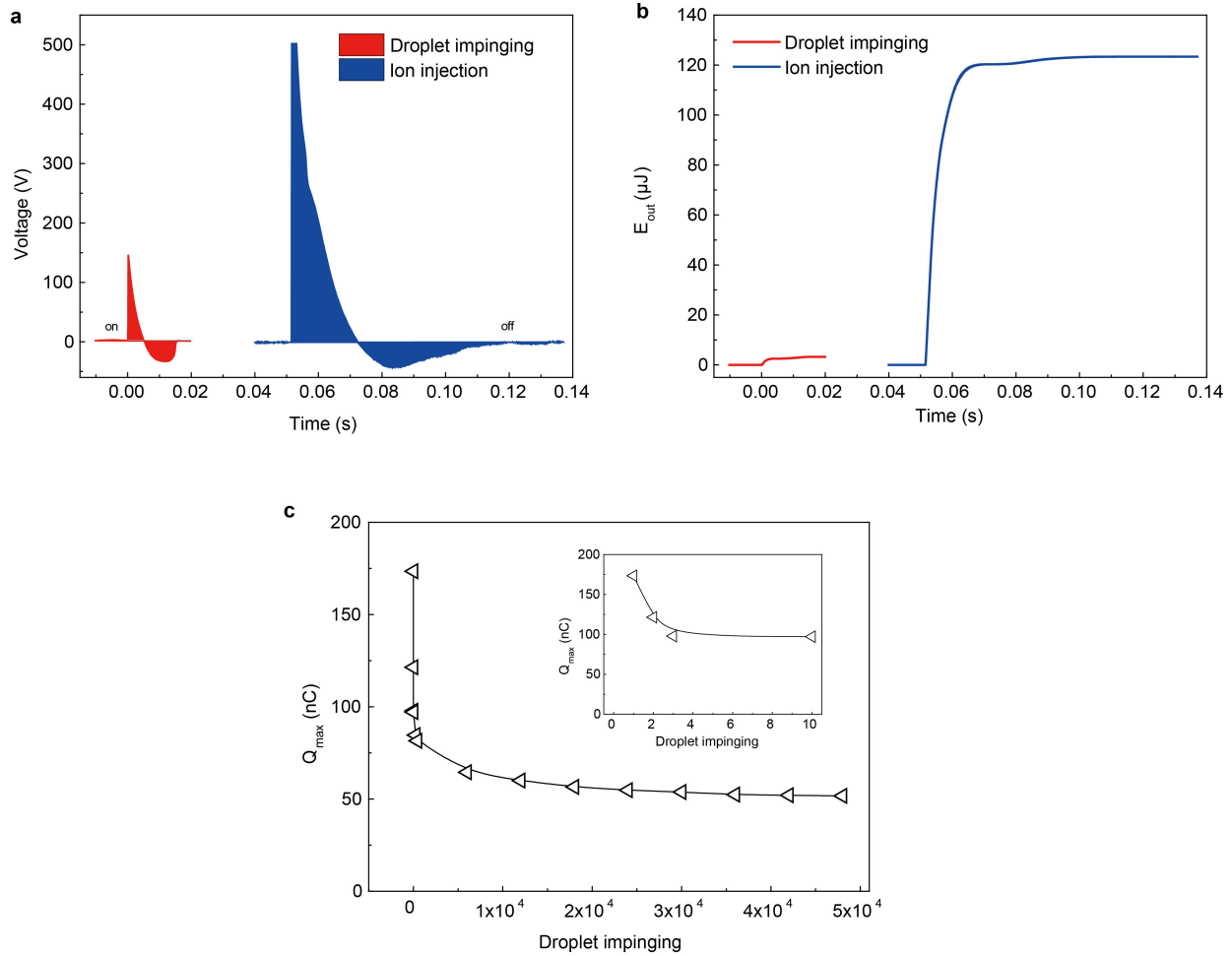
Extended Data Fig. 2 | Surface charging by continuous droplet impinging, and characterization of the output charge and power of the DEG. **a**, Variation in the output voltage as a function of the number of individual impinging droplets. In this case, the surface was not precharged. The output results purely from charge generation and transfer during droplet impinging. **b**, The output charge measured from the DEG (in red; the frequency of impinging

droplets is 4.2 Hz and the total number of droplets is about 42) is roughly 49.8 nC, which is around 640.1 times higher than that of the control device (in black; the frequency of impinging droplets is 1.0 Hz, and the total number of droplets is 9). **c**, When the load resistance increases from 1 k Ω to 100 M Ω , the output current decreases from 250.0 μA to 2.0 μA . When the load resistance is 332.0 k Ω , the peak output power density is 50.1 W m^{-2} .



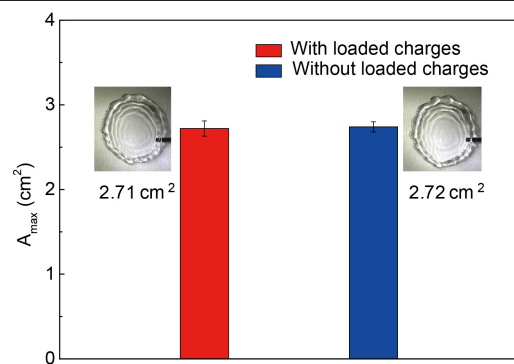
Extended Data Fig. 3 | Control experiment based on a triboelectric nanogenerator. **a**, Optical image showing the as-fabricated control device. The structure of this control device is similar to that of the DEG, but without an aluminium electrode. **b**, Diagram showing its detailed working mechanism. **i**, Before the droplet contacts the PTFE, the amount of (positive) charges on the ITO is the same as the (negative) charges on the PTFE, owing to electrostatic induction. Thus, there is no current flow from ITO to the ground. **ii**, When a water droplet contacts the PTFE surface, the droplet becomes positively charged while the PTFE becomes more negatively charged as a result of contact electrification. **iii**, When the positively charged droplet leaves, it causes the ITO electrode with positive charges to be unable to screen the more negatively

charged PTFE. **iv**, Accordingly, a flow of current (I) between the ground and ITO electrode is induced by electrostatic induction. **c**, Variation in voltage output from the control device as a result of continuous droplet impinging. The inset shows the time-dependent variation in voltage from a single droplet. The frequency of impinging droplets is set at 1.0 Hz, with a total of nine droplets. **d**, In each test of the control device, the amount of transferred charge (in green) is identical to the charge carried by the departing droplets (in blue), showing that electricity generation from the control device indeed originates from contact electrification. The frequency of impinging droplets is 1.0 Hz, with a total of nine droplets.

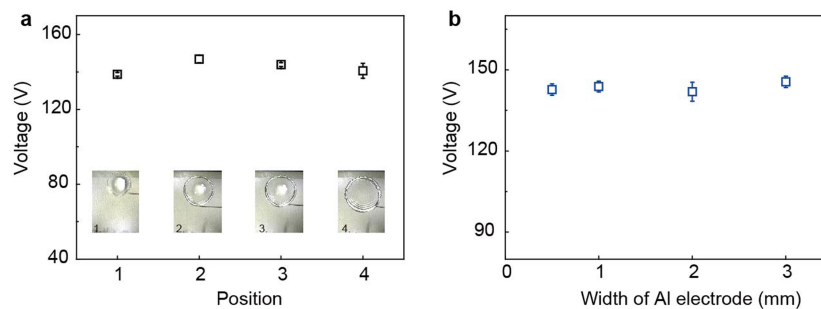


Extended Data Fig. 4 | Enhanced electrical output using the ion-injection method. **a**, Comparison of the output voltage generated from a single droplet impinging on a DEG that was precharged by droplet impinging (in red) or by ion injection (in blue). **b**, Comparison of the amount of electrical energy (E_{out}) generated from a single droplet impinging on a DEG charged by droplet impinging (in red) or ion injection (in blue). The instantaneous peak density can be enhanced further by increasing the surface charge on the PTFE film

through ion injection, using a commercial antistatic gun (Zerostat3, Milty) to inject various ions, including CO_3^{3-} , NO_3^{3-} , NO_2^{2-} , O_3^{3-} and O_2^{2-} , from a vertical distance of roughly 5.0 cm. **c**, Variation in the measured maximum charge, Q_{max} , with droplet impinging on a PTFE surface that was precharged using ion injection. Q_{max} decays rapidly and finally reaches a stable value of roughly 49.8 nC. The inset shows the Q_{max} for the first four droplets.

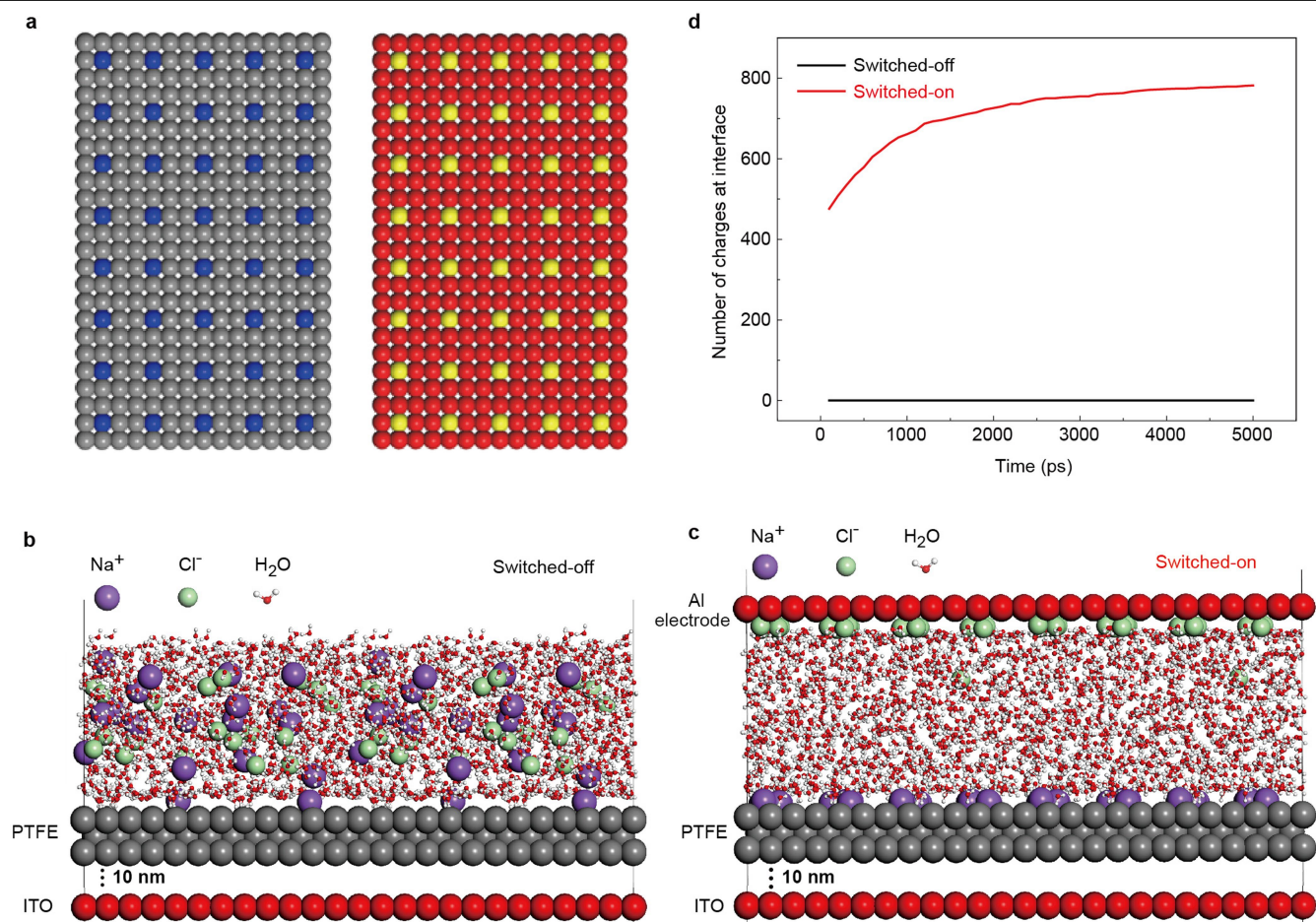


Extended Data Fig. 5 | Effect of surface charge on the maximum spreading area, A_{\max} , of a droplet. Data are means \pm s.e.m. For each mean, the total number of measurements is ten.



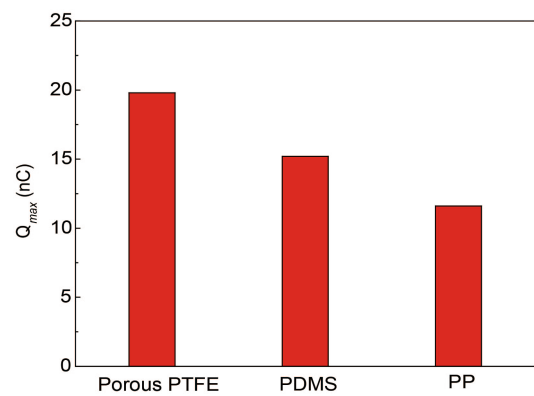
Extended Data Fig. 6 | Effect of the spatial location and width of the aluminium electrode on electricity generation. **a**, The spatial location of the aluminium electrode was changed, keeping the impact location fixed. In this way, the spacing between the droplet centre and the electrode can be tailored. Insets marked with 1, 2, 3, 4 refer to the four different locations of the aluminium electrode on the PTFE surface. The results show that regardless of the electrode location, the output voltage is constant, suggesting that

electricity generation is not sensitive to electrode location. Data are means \pm s.e.m. For each mean, the total number of measurements is ten. **b**, The output voltage does not depend on the size of the aluminium electrode. This makes sense because the source of electricity generation is the electrostatically induced charge on the ITO, rather than on the aluminium. Data are means \pm s.e.m. For each mean, the total number of measurements is ten.

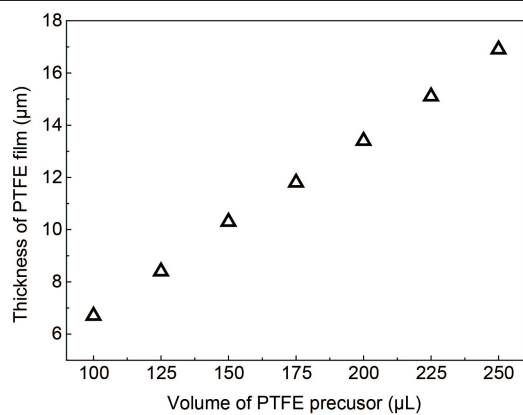


Extended Data Fig. 7 | Molecular-dynamics simulation. **a**, In the molecular-dynamics simulation, negative (blue) and positive (yellow) charges are fixed on atomic layers of PTFE (grey) and ITO (red), respectively. **b**, Molecular-dynamics simulation showing the distribution of mobile charges (Na^+ and Cl^-) inside the water and on the PTFE surface in switched-off mode (that is, without an

aluminium electrode, although the negative and positive charges on PTFE and ITO are turned on). **c**, Molecular-dynamics simulation showing the distribution of charges inside the water and on the PTFE surface in switched-on mode. **d**, Comparison of the number of mobile charges transferred to the water/solid interface in switched-on and switched-off modes.



Extended Data Fig. 8 | Control devices made of porous PTFE, PDMS and PP. Comparison of the maximum stable surface charge, Q_{max} , on control devices made of porous PTFE, PDMS and polypropylene (PP) after continuous droplet impact under a relative humidity of 65.0%; all of these charges are much smaller than that of our DEG surface.



Extended Data Fig. 9 | Thickness of PTFE film as a function of the volume of PTFE precursor. The thickness of the PTFE film increases linearly from 6.7 μm to 16.9 μm as the volume of PTFE precursor increases from 100.0 μl to 250.0 μl .

Closed-loop optimization of fast-charging protocols for batteries with machine learning

<https://doi.org/10.1038/s41586-020-1994-5>

Received: 6 August 2019

Accepted: 19 December 2019

Published online: 19 February 2020

 Check for updates

Peter M. Attia^{1,7}, Aditya Grover^{2,7}, Norman Jin¹, Kristen A. Severson³, Todor M. Markov², Yang-Hung Liao¹, Michael H. Chen¹, Bryan Cheong^{1,2}, Nicholas Perkins¹, Zi Yang¹, Patrick K. Herring⁴, Muratahan Aykol⁴, Stephen J. Harris^{1,5}, Richard D. Braatz^{3,✉}, Stefano Ermon^{2,✉} & William C. Chueh^{1,6,✉}

Simultaneously optimizing many design parameters in time-consuming experiments causes bottlenecks in a broad range of scientific and engineering disciplines^{1,2}. One such example is process and control optimization for lithium-ion batteries during materials selection, cell manufacturing and operation. A typical objective is to maximize battery lifetime; however, conducting even a single experiment to evaluate lifetime can take months to years^{3–5}. Furthermore, both large parameter spaces and high sampling variability^{3,6,7} necessitate a large number of experiments. Hence, the key challenge is to reduce both the number and the duration of the experiments required. Here we develop and demonstrate a machine learning methodology to efficiently optimize a parameter space specifying the current and voltage profiles of six-step, ten-minute fast-charging protocols for maximizing battery cycle life, which can alleviate range anxiety for electric-vehicle users^{8,9}. We combine two key elements to reduce the optimization cost: an early-prediction model⁵, which reduces the time per experiment by predicting the final cycle life using data from the first few cycles, and a Bayesian optimization algorithm^{10,11}, which reduces the number of experiments by balancing exploration and exploitation to efficiently probe the parameter space of charging protocols. Using this methodology, we rapidly identify high-cycle-life charging protocols among 224 candidates in 16 days (compared with over 500 days using exhaustive search without early prediction), and subsequently validate the accuracy and efficiency of our optimization approach. Our closed-loop methodology automatically incorporates feedback from past experiments to inform future decisions and can be generalized to other applications in battery design and, more broadly, other scientific domains that involve time-intensive experiments and multi-dimensional design spaces.

Optimal experimental design (OED) approaches are widely used to reduce the cost of experimental optimization. These approaches often involve a closed-loop pipeline where feedback from completed experiments informs subsequent experimental decisions, balancing the competing demands of exploration—that is, testing regions of the experimental parameter space with high uncertainty—and exploitation—that is, testing promising regions based on the results of the completed experiments. Adaptive OED algorithms have been successfully applied to physical science domains, such as materials science^{1,2,12–14}, chemistry^{15,16}, biology¹⁷ and drug discovery¹⁸, as well as to computer science domains, such as hyperparameter optimization for machine learning^{19,20}. However, while a closed-loop approach is designed to

minimize the number of experiments required for optimizing across a multi-dimensional parameter space, the time (and cost) per experiment may remain high, as is the case for lithium-ion batteries. Therefore, an OED approach should account for both the number of experiments and the cost per experiment. Multi-fidelity optimization approaches have been developed to learn from both inexpensive, noisy signals and expensive, accurate signals. For example, in hyperparameter optimization for machine learning algorithms, several low-fidelity signals for predicting the final performance of an algorithmic configuration (for example, extrapolated learning curves^{19,20}, rapid testing on a subset of the full training dataset²¹) are used in tandem with more complete configuration evaluations^{22,23}. For lithium-ion batteries, classical

¹Department of Materials Science and Engineering, Stanford University, Stanford, CA, USA. ²Department of Computer Science, Stanford University, Stanford, CA, USA. ³Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. ⁴Toyota Research Institute, Los Altos, CA, USA. ⁵Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. ⁶Applied Energy Division, SLAC National Accelerator Laboratory, Menlo Park, CA, USA. ⁷These authors contributed equally: Peter M. Attia, Aditya Grover. ✉e-mail: braatz@mit.edu; ermon@cs.stanford.edu; wchueh@stanford.edu

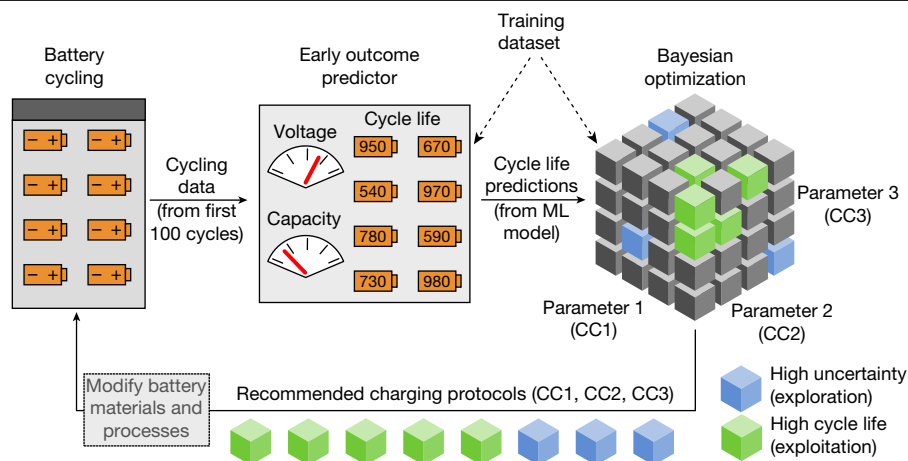


Fig. 1 | Schematic of our CLO system. First, batteries are tested. The cycling data from the first 100 cycles (specifically, electrochemical measurements such as voltage and capacity) are used as input for an early outcome prediction of cycle life. These cycle life predictions from a machine learning (ML) model are subsequently sent to a BO algorithm, which recommends the next protocols to test by balancing the competing demands of exploration (testing protocols with high uncertainty in estimated cycle life) and exploitation

(testing protocols with high estimated cycle life). This process iterates until the testing budget is exhausted. In this approach, early prediction reduces the number of cycles required per tested battery, while optimal experimental design reduces the number of experiments required. A small training dataset of batteries cycled to failure is used both to train the early outcome predictor and to set BO hyperparameters. In future work, design of battery materials and processes could also be integrated into this closed-loop system.

methods such as factorial design that use predetermined heuristics to select experiments have been applied^{24–26}, but the design and use of low-fidelity signals is challenging and unexplored. These previously considered approaches do not discover and exploit the patterns present in the parameter space for efficient optimization, nor do they address the issue of time per experiment.

In this work, we develop a closed-loop optimization (CLO) system with early outcome prediction for efficient optimization over large parameter spaces with expensive experiments and high sampling variability. We employ this system to experimentally optimize fast-charging protocols for lithium-ion batteries; reducing charging times to approach gasoline refuelling time is critical to reduce range anxiety for electric vehicles^{8,9} but often comes at the expense of battery lifetime. Specifically, we optimize over a parameter space consisting of 224 unique six-step, ten-minute fast-charging protocols (that is, how current and voltage are controlled during charging) to find charging protocols with high cycle life (defined as the battery capacity falling to 80% of its nominal value). Our system uses two key elements to reduce the optimization cost (Extended Data Fig. 1). First, we reduce the time per experiment by using machine learning to predict the outcome of the experiment based on data from early cycles, well before the batteries reach the end of life⁵. Second, we reduce the number of experiments by using a Bayesian optimization (BO) algorithm to balance the exploration–exploitation tradeoff in choosing the next round of experiments^{10,11}. Testing a single battery to failure under our fast-charging conditions requires approximately 40 days, meaning that when 48 experiments are performed in parallel, assessing all 224 charging protocols with triplicate measurements takes approximately 560 days. Here, using CLO with early outcome prediction, only 16 days were required to confidently identify protocols with high cycle lives (48 parallel experiments). In a subsequent validation study, we find that CLO ranks these protocols by lifetime accurately (Kendall rank correlation coefficient, 0.83) and efficiently (15 times less time than a baseline ‘brute-force’ approach that uses random search without early prediction). Furthermore, we find that the charging protocols identified as optimal by CLO with early prediction outperform existing fast-charging protocols designed to avoid lithium plating (a common fast-charging degradation mode), the approach suggested by conventional battery wisdom^{4,8,9,26}. This work highlights the utility of combining CLO with inexpensive early outcome predictors to accelerate scientific discovery.

CLO with early outcome prediction is depicted schematically in Fig. 1. The system consists of three components: parallel battery cycling, an early predictor for cycle life and a BO algorithm. At each sequential round, we iterate over these three components. The first component is a multi-channel battery cycler; the cycler used in this work tests 48 batteries simultaneously. Before starting CLO, the charging protocols for the first round of 48 batteries are chosen at random (without replacement) from the complete set of 224 unique multi-step protocols (Methods). Each battery undergoes repeated charging and discharging for 100 cycles (about 4 days; average predicted cycle life 905 cycles), beyond which the experiments are terminated.

These cycling data are then fed as input to the early outcome predictor, which estimates the final cycle lives of the batteries given data from the first 100 cycles. The early predictor is a linear model trained via elastic net regression²⁷ on features extracted from the charging data of the first 100 cycles (Supplementary Table 1), similar to that presented in Severson et al.⁵. Predictive features include transformations of both differences between voltage curves and discharge capacity fade trends. To train the early predictor, we require a training dataset of batteries cycled to failure. Here, we used a pre-existing dataset of 41 batteries cycled to failure (cross-validation root-mean-square error, 80.4 cycles; see Methods and Supplementary Discussion 1). Whereas obtaining this dataset itself requires running full cycling experiments for a small training set of batteries (the cost we are trying to offset), this one-time cost could be avoided if pretrained predictors or previously collected datasets are available. If unavailable, we pay an upfront cost in collecting this dataset; this dataset could also be used for warm-starting the BO algorithm. The size of the dataset collected should best tradeoff the upfront cost in acquiring the dataset to train an accurate model with the anticipated reduction in experimentation requirements for CLO.

Finally, these predicted cycle lives from early-cycle data are fed into the BO algorithm (Methods and Supplementary Discussion 2), which recommends the next round of 48 charging protocols that best balance the exploration–exploitation tradeoff. This algorithm (Methods and Supplementary Discussion 2) builds on the prior work of Hoffman et al.¹⁰ and Grover et al.¹¹. The algorithm maintains an estimate of both the average cycle life and the uncertainty bounds for each protocol; these estimates are initially equal for all protocols and are refined as additional data are collected. Crucially, to reduce the total optimization cost, our algorithm performs these updates using estimates from

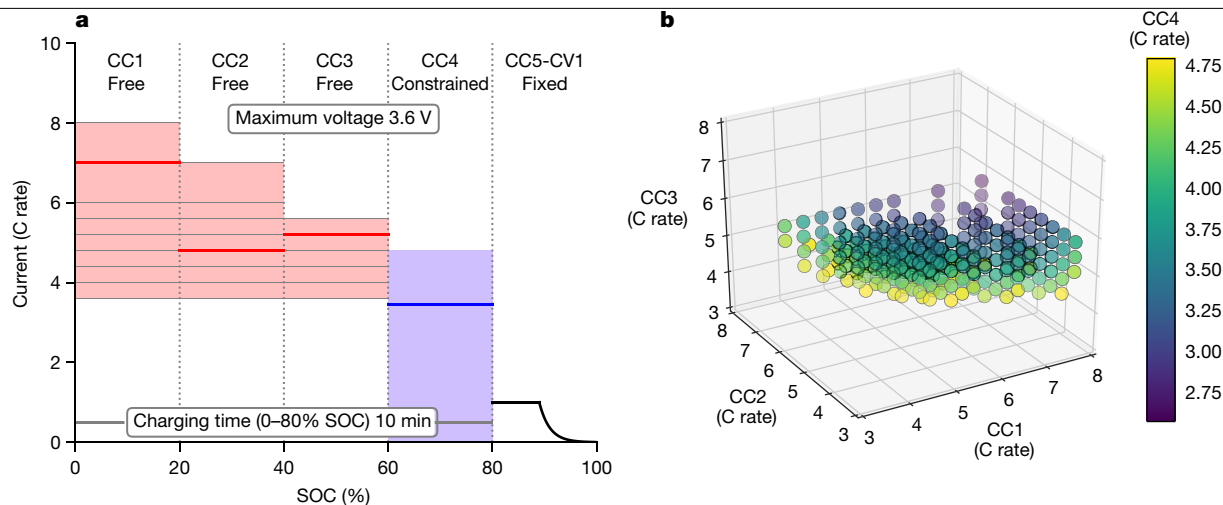


Fig. 2 | Structure of our six-step, ten-minute fast-charging protocols.

Currents are defined as dimensionless C rates; here, 1C is 1.1 A, or the current required to fully (dis)charge the nominal capacity (1.1 Ah) in 1 h. **a**, Current versus SOC for an example charging protocol, 7.0C–4.8C–5.2C–3.45C (bold lines). Each charging protocol is defined by five constant current (CC) steps followed by one constant voltage (CV) step. The last two steps (CC5 and CV1) are identical for all charging protocols. We optimize over the first four constant-current steps, denoted CC1, CC2, CC3 and CC4. Each of these steps comprises a 20% SOC window, such that CC1 ranges from 0% to 20% SOC, CC2

ranges from 20% to 40% SOC, and so on. CC4 is constrained by specifying that all protocols charge in the same total time (10 min) from 0% to 80% SOC. Thus, our parameter space consists of unique combinations of the three free parameters CC1, CC2 and CC3. For each step, we specify a range of acceptable values; the upper limit is monotonically decreasing with increasing SOC to avoid the upper cutoff potential (3.6 V for all steps). **b**, CC4 (colour scale) as a function of CC1, CC2 and CC3 (on the x, y and z axes, respectively). Each point represents a unique charging protocol.

the early outcome predictor instead of using the actual cycle lives. The mean and uncertainty estimates for the cycle lives are obtained via a Gaussian process (Methods), which has a smoothing effect and allows for updating the cycle life estimates of untested protocols with the predictions from related protocols. The closed-loop process repeats until the optimization budget, in our case 192 batteries tested (100 cycles each), is exhausted.

Our objective is to find the charging protocol which maximizes the expected battery cycle life for a fixed charging time (ten minutes) and state-of-charge (SOC) range (0 to 80%). The design space of our 224 six-step extreme fast-charging protocols is presented in Fig. 2a. Multi-step charging protocols, in which a series of different constant-current steps are applied within a single charge, are considered advantageous over single-step charging for maximizing cycle life during fast charging^{4,8}, though the optimal combination remains unclear. As shown in Fig. 2b, each protocol is specified by three independent parameters (CC1, CC2 and CC3); each parameter is a current applied over a fixed SOC range (0–20%, 20–40% and 40–60%, respectively). A fourth parameter, CC4, is dependent on CC1, CC2, CC3 and the charging time. Given constraints on the current values (Methods), a total of 224 charging protocols are permitted. We test commercial lithium iron phosphate (LFP)/graphite cylindrical batteries (A123 Systems) in a convective environmental chamber (30 °C ambient temperature). A maximum voltage of 3.6 V is imposed. These batteries are designed to fast-charge in 17 min (rate testing data are presented in Extended Data Fig. 2). The cycle life decreases dramatically with faster charging time^{4,5}, motivating this optimization. Since the LFP positive electrode is generally considered to be stable^{4,5}, we select this battery chemistry to isolate the effects of extreme fast charging on graphite, which is universally employed in lithium-ion batteries.

In all, we ran four CLO rounds sequentially, consisting of 185 batteries in total (excluding seven batteries; see Methods). Using early prediction, each CLO round requires four days to complete 100 cycles, resulting in a total testing time of sixteen days—a major reduction from the 560 days required to test each charging protocol to failure three times. Figure 3 presents the predictions and selected protocols (Fig. 3a), as well as the evolution of cycle life estimates over the parameter space

as the optimization progresses (Fig. 3a). Initially, the estimated cycle lives for all protocols are equal. After two rounds, the overall structure of the parameter space (that is, the dependence of cycle life on charging protocol parameters CC1, CC2 and CC3) emerges, and a prominent region with high cycle life protocols has been identified. The confidence of CLO in this high-performing region is further improved from round 2 to round 4, but overall the cycle life estimates do not change substantially (Extended Data Fig. 3). By learning and exploiting the structure of the parameter space, we avoid evaluating charging protocols with low estimated cycle life and concentrate more resources on the high-performing region (Extended Data Figs. 3–5). Specifically, 117 of 224 protocols are never tested (Fig. 3c); we spend 67% of the batteries testing 21% of the protocols (0.83 batteries per protocol on average). CLO repeatedly tests several protocols with high estimated cycle life to decrease uncertainties due to manufacturing variability and the error introduced by early outcome prediction. The uncertainty is expressed as the prediction intervals of the posterior predictive distribution over cycle life (Extended Data Figs. 3g, 4, 5).

To the best of our knowledge, this work presents the largest known map of cycle life as a function of charging conditions (Extended Data Fig. 5). This dataset can be used to validate physics-based models of battery degradation. Most fast-charging protocols proposed in the battery literature suggest that current steps decreasing monotonically as a function of SOC are optimal to avoid lithium plating on graphite, a well-accepted degradation mode during fast charging^{4,8,9,26}. In contrast, the protocols identified as optimal by CLO (for example, Fig. 3d) are generally similar to single-step constant-current charging (that is, $CC_1 \approx CC_2 \approx CC_3 \approx CC_4$). Specifically, of the 75 protocols with the highest estimated cycle lives, only ten are monotonically decreasing (that is, $CC_i \geq CC_{i+1}$ for all i) and two are strictly decreasing (that is, $CC_i > CC_{i+1}$). We speculate that minimizing parasitic reactions caused by heat generation may be the operative optimization strategy for these cells, as opposed to minimizing the propensity for lithium plating (Supplementary Discussion 3). While the optimal protocol for a new scenario would depend on the selected charge time, SOC window, temperature control conditions and battery chemistry, this unexpected result highlights the need for data-driven approaches for optimizing fast charging.

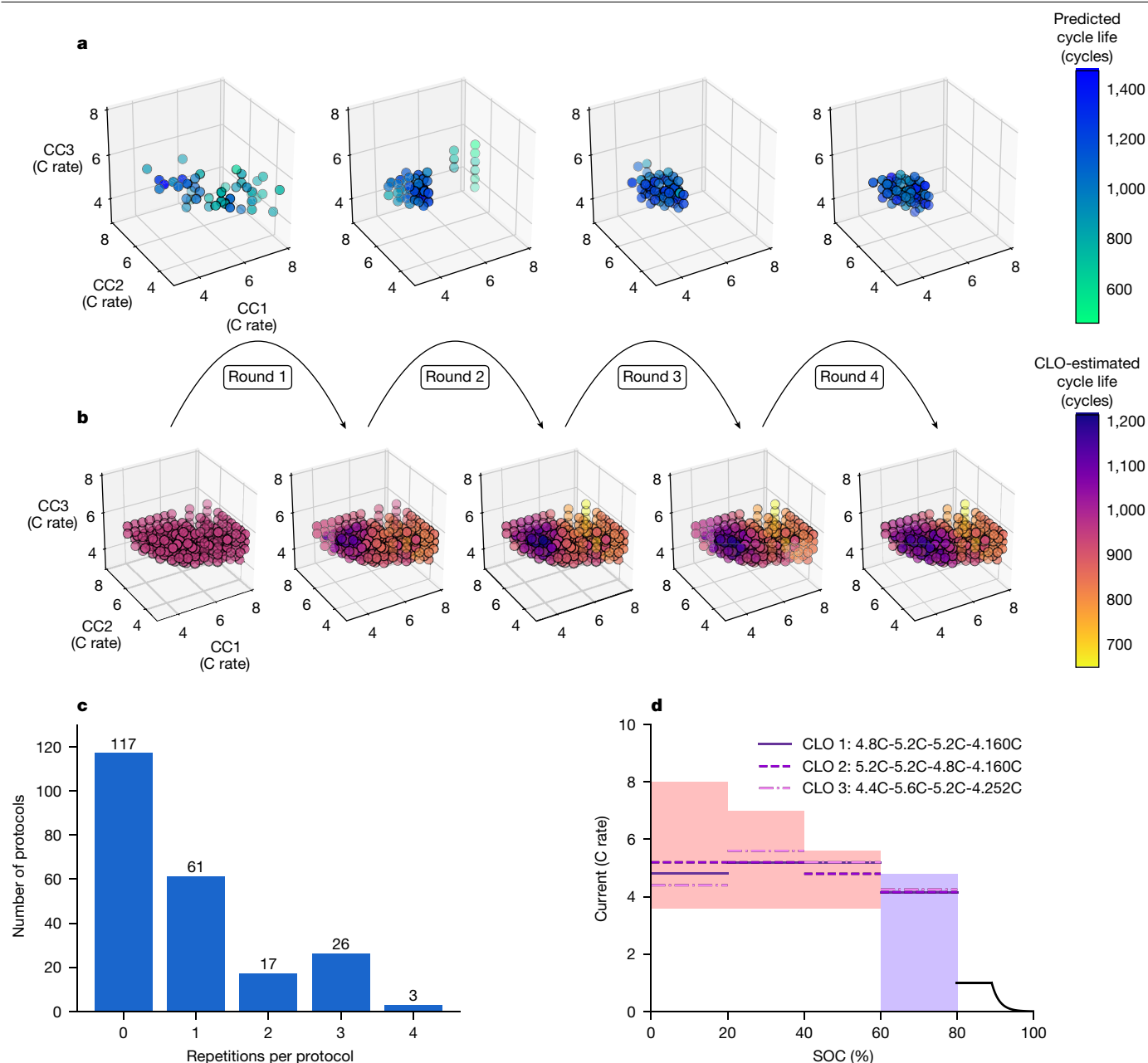


Fig. 3 | Results of closed-loop experiments. **a**, Early cycle life predictions per round. The tested charging protocols and the resulting predictions are plotted for rounds 1–4. Each point represents a charging protocol, defined by CC1, CC2 and CC3 (the x, y and z axes, respectively). The colour scale represents cycle life predictions from the early outcome prediction model. The charging protocols in the first round of testing are randomly selected. As the BO algorithm shifts from exploration to exploitation, the charging protocols selected for testing by the closed loop in subsequent rounds fall primarily into the high-performing region. **b**, Evolution of the parameter space per round. The colour scale represents cycle life, as estimated by the BO algorithm. The initial cycle life

estimates are equivalent for all protocols; as more predictions are generated, the BO algorithm updates its cycle life estimates. The CLO-estimated mean cycle lives after four rounds for all fast-charging protocols in the parameter space are also presented in Extended Data Fig. 5 and Supplementary Table 3. **c**, Distribution of the number of repetitions for each charging protocol (excluding failed batteries). Only 46 of 224 protocols (21%) are tested multiple times. **d**, Current versus SOC for the top three fast-charging protocols, as estimated by CLO. CC1–CC4 are displayed in the legend. All three protocols have relatively uniform charging (that is, $CC1 \approx CC2 \approx CC3 \approx CC4$).

We validate the performance of CLO with early prediction on a subset of nine extreme fast-charging protocols. For each of these protocols, we cycle five batteries each to failure and use the sample average of the final cycle lives as an estimate of the true lifetimes. We use this validation study to (1) confirm that CLO is able to correctly rank protocols based on cycle life, (2) compare the cycle lives of protocols recommended by CLO to protocols inspired by the battery literature and (3) compare the performance of CLO to baseline ablation approaches for experimental design. The charging protocols used in validation,

some of which are inspired by existing battery fast-charging literature (see Methods), span the range of estimated cycle lives (Extended Data Fig. 6 and Extended Data Table 1). We adjust the voltage limits and charging times of these literature protocols to match our protocols, while maintaining similar current ratios as a function of SOC. Whereas the literature protocols used in these validation experiments are generally designed for batteries with high-voltage positive electrode chemistries, fast-charging optimization strategies generally focus on the graphitic negative electrode^{4,8}. For these nine protocols, we validate

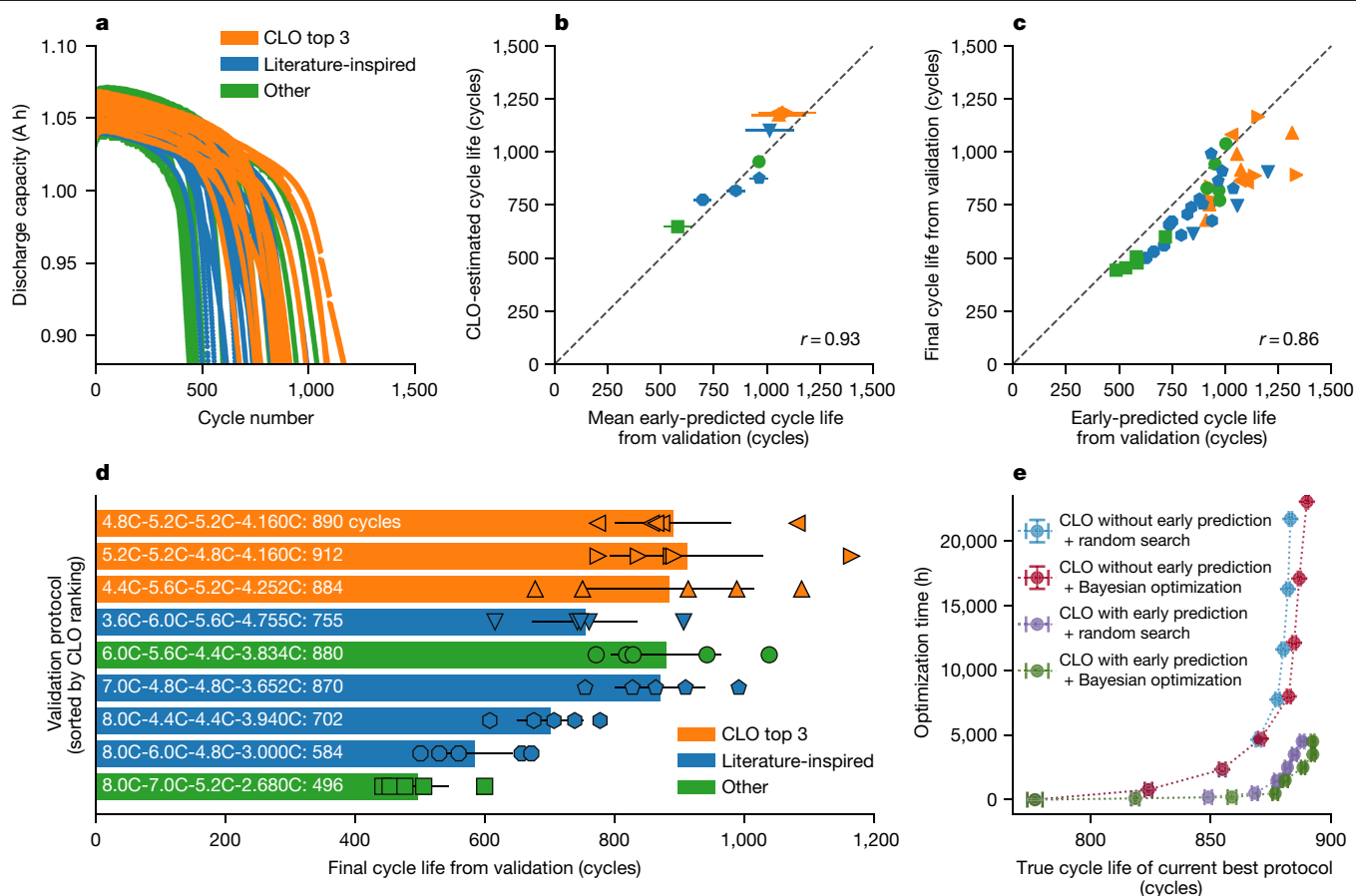


Fig. 4 | Results of validation experiment. **a**, Discharge capacity versus cycle number for all batteries in the validation experiment. The nine validation protocols include the top three protocols as estimated by CLO ('CLO top 3'), four protocols inspired by the battery literature^{39–44} ('Literature-inspired') and two protocols selected to obtain a representative sampling from the distribution of CLO-estimated cycle lives among the validation protocols ('Other'). **b**, Comparison of early-predicted cycle lives from validation to closed-loop estimates, averaged on a protocol basis. Each ten-minute charging protocol is tested with five batteries. Error bars represent the 95% confidence intervals. **c**, Observed versus early-predicted cycle life for the validation experiment. Although our early predictor tends to overestimate cycle life,

probably owing to calendar ageing effects (Supplementary Discussion 4), the trend is correctly captured (Pearson correlation coefficient $r = 0.86$). **d**, Final cycle lives from validation, sorted by CLO ranking. The length of each bar and the annotations represents the mean final cycle life from validation per protocol. Error bars represent the 95% confidence intervals. **e**, Ablation study of various optimization approaches using the protocols and data in the validation set (Methods). Error bars represent the 95% confidence intervals ($n = 2,000$). With contributions from both early prediction and Bayesian optimization, CLO can rapidly identify high-performing charging protocols. The gains from Bayesian optimization are larger when resources are constrained (Extended Data Fig. 8).

the 'CLO-estimated' cycle lives against the sample average of the five final cycle lives.

The validation results are presented in Fig. 4. The discharge capacity fade curves (Fig. 4a) exhibit the nonlinear decay typical of fast charging^{5,7}. If we apply our early-prediction model to the batteries in the validation experiment, these early predictions (averaged over each protocol) match the CLO-estimated mean cycle lives well (Pearson correlation coefficient $r = 0.93$; Fig. 4b). This result validates the performance of the BO component of CLO in particular, since the CLO-estimated cycle lives were inferred from early predictions. However, our early-prediction model exhibits some bias (Fig. 4c), probably owing to calendar ageing effects from different battery storage times²⁸ (Supplementary Table 2 and Supplementary Discussion 4). Despite this bias in our predictive model, we generally capture the ranking well (Kendall rank correlation coefficient, 0.83; Fig. 4d and Extended Data Fig. 7). At the same time, we note that the final cycle lives for the top-ranked protocols are similar. Furthermore, the optimal protocols identified by CLO outperform protocols inspired by previously published fast-charging protocols (895 versus 728 cycles on average; Extended Data Fig. 6 and Extended Data Table 1). This result suggests that the efficiency of our approach does not come at the expense of accuracy.

Our method greatly reduces the optimization time required compared to baseline optimization approaches (Fig. 4e). For instance, a procedure that does not use early outcome prediction and simply selects protocols randomly to test begins to saturate at a competitive performance level after about 7,700 battery-hours of testing. To achieve a similar level of performance, CLO with both early outcome prediction and the BO algorithm requires only 500 battery-hours of testing. For this small-scale validation experiment, we observe that the early-prediction component of CLO greatly reduces the time per experiment. Here, random selection is equivalent to a pure exploration strategy and can achieve a performance similar to the BO-based approaches for smaller experimental budgets. In later stages, random selection is eventually outperformed by BO-based approaches, which exploit the structure across the protocols and focus on reducing the uncertainty in the promising regions of the parameter space. Although these results are specific to this validation study, we observe similar or larger gains in simulations when fewer batteries or fewer parallel experiments (relative to the size of the parameter space) are available (Extended Data Fig. 8). The relative gains from BO over random selection are largest with minimal resources.

Finally, we compare our early predictor with other low-fidelity predictors proposed in state-of-the-art multi-fidelity optimization algorithms

in the literature^{19,20}, and find that our approach outperforms these algorithms (Supplementary Discussion 2 and Supplementary Table 4). The generic early-prediction models in these previous works fit composites of parametric functions to the capacity fade curves, while our model uses additional features recorded at every cycle (for example, voltage). This result highlights the value of designing predictive models for the target application in multi-fidelity optimization.

In summary, we have successfully accelerated the optimization of extreme fast charging for lithium-ion batteries using CLO with early outcome prediction. This method could extend to other fast-charging design spaces, such as pulsed^{26,28} and constant-power⁸ charging, as well as to other objectives, such as slower charging and calendar ageing. Additionally, this work opens up new applications for battery optimization, such as formation²⁹, adaptive cycling³⁰ and parameter estimation for battery management system models³¹. Furthermore, provided that a suitable early outcome predictor exists, this method could also be applied to optimize other aspects of battery development, such as electrode materials and electrolyte chemistries. Beyond batteries, our CLO approach combining black-box optimization with early outcome prediction can be extended to efficiently optimize other physical^{1,2,18} and computational^{22,32} multi-dimensional parameter spaces that involve time-intensive experimentation, illustrating the power of data-driven methods to accelerate the pace of scientific discovery.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-1994-5>.

1. Tabor, D. P. et al. Accelerating the discovery of materials for clean energy in the era of smart automation. *Nat. Rev. Mater.* **3**, 5–20 (2018).
2. Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O. & Walsh, A. Machine learning for molecular and materials science. *Nature* **559**, 547–555 (2018).
3. Baumhöfer, T., Brühl, M., Rothgang, S. & Sauer, D. U. Production caused variation in capacity aging trend and correlation to initial cell performance. *J. Power Sources* **247**, 332–338 (2014).
4. Keil, P. & Jossen, A. Charging protocols for lithium-ion batteries and their impact on cycle life—an experimental study with different 18650 high-power cells. *J. Energy Storage* **6**, 125–141 (2016).
5. Severson, K. A. et al. Data-driven prediction of battery cycle life before capacity degradation. *Nat. Energy* **4**, 383–391 (2019).
6. Schuster, S. F., Brand, M. J., Berg, P., Gleissenberger, M. & Jossen, A. Lithium-ion cell-to-cell variation during battery electric vehicle operation. *J. Power Sources* **297**, 242–251 (2015).
7. Harris, S. J., Harris, D. J. & Li, C. Failure statistics for commercial lithium ion batteries: a study of 24 pouch cells. *J. Power Sources* **342**, 589–597 (2017).
8. Ahmed, S. et al. Enabling fast charging—a battery technology gap assessment. *J. Power Sources* **367**, 250–262 (2017).
9. Liu, Y., Zhu, Y. & Cui, Y. Challenges and opportunities towards fast-charging battery materials. *Nat. Energy* **4**, 540–550 (2019).
10. Hoffman, M. W., Shahriari, B. & de Freitas, N. On correlation and budget constraints in model-based bandit optimization with application to automatic machine learning. In *Proc. 17th Int. Conf. on Artificial Intelligence and Statistics (AISTATS)* Vol. 33, 365–374 (Proceedings of Machine Learning Research, 2014); <http://proceedings.mlr.press/v33/hoffman14.html>.
11. Grover, A. et al. Best arm identification in multi-armed bandits with delayed feedback. In *Proc. 21st Int. Conf. on Artificial Intelligence and Statistics (AISTATS)* Vol. 84, 833–842 (Proceedings of Machine Learning Research, 2018); <http://proceedings.mlr.press/v84/grover18b.html>.
12. Nikolaev, P. et al. Autonomy in materials research: a case study in carbon nanotube growth. *npj Comput. Mater.* **2**, 16031 (2016).
13. Ling, J., Hutchinson, M., Antono, E., Paradiso, S. & Meredig, B. High-dimensional materials and process optimization using data-driven experimental design with well-calibrated uncertainty estimates. *Integr. Mater. Manuf. Innov.* **6**, 207–217 (2017).
14. Balachandran, P. V., Kowalski, B., Sehirioglu, A. & Lookman, T. Experimental search for high-temperature ferroelectric perovskites guided by two-step machine learning. *Nat. Commun.* **9**, 1668 (2018).
15. Bédard, A.-C. et al. Reconfigurable system for automated optimization of diverse chemical reactions. *Science* **361**, 1220–1225 (2018).
16. Granda, J. M., Donina, L., Dragone, V., Long, D.-L. & Cronin, L. Controlling an organic synthesis robot with machine learning to search for new reactivity. *Nature* **559**, 377–381 (2018).
17. King, R. D. et al. The automation of science. *Science* **324**, 85–89 (2009).
18. Schneider, G. Automating drug discovery. *Nat. Rev. Drug Discov.* **17**, 97–113 (2018).
19. Domhan, T., Springenberg, J. T. & Hutter, F. Speeding up automatic hyperparameter optimization of deep neural networks by extrapolation of learning curves. In *Proc. 24th Int. Conf. on Artificial Intelligence* 3460–3468 (AAAI Press, 2015).
20. Klein, A., Falkner, S., Springenberg, J. T. & Hutter, F. Learning curve prediction with Bayesian neural networks. In *Proc. 2017 Int. Conf. on Learning Representations* 1–16 (2017); <https://openreview.net/forum?id=S11KBVclx>.
21. Petrak, J. *Fast Subsampling Performance Estimates for Classification Algorithm Selection*. Technical Report TR-2000-07, 3–14 (Austrian Research Institute for Artificial Intelligence, 2000); <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.28.3305&rep=rep1&type=pdf>.
22. Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A. & Talwalkar, A. Hyperband: a novel bandit-based approach to hyperparameter optimization. *J. Mach. Learn. Res.* **18**, 1–52 (2018).
23. Hutter, F., Hoos, H. H. & Leyton-Brown, K. Sequential model-based optimization for general algorithm configuration. In *Proc. 5th Int. Conf. on Learning and Intelligent Optimization* 507–523 (Springer, 2011).
24. Luo, Y., Liu, Y. & Wang, S. Search for an optimal multistage charging pattern for lithium-ion batteries using the Taguchi approach. In *Region 10 Conf. (TENCON 2009)* 1–5, <https://doi.org/10.1109/TENCON.2009.5395823> (IEEE, 2009).
25. Liu, Y., Hsieh, C. & Luo, Y. Search for an optimal five-step charging pattern for Li-ion batteries using consecutive orthogonal arrays. *IEEE Trans. Energy. Convers.* **26**, 654–661 (2011).
26. Schindler, S., Bauer, M., Cheetamun, H. & Danzer, M. A. Fast charging of lithium-ion cells: identification of aging-minimal current profiles using a design of experiment approach and a mechanistic degradation analysis. *J. Energy Storage* **19**, 364–378 (2018).
27. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B* **67**, 301–320 (2005).
28. Keil, P. et al. Calendar aging of lithium-ion batteries. I. Impact of the graphite anode on capacity fade. *J. Electrochem. Soc.* **163**, A1872–A1880 (2016).
29. Wood, D. L., Li, J. & Daniel, C. Prospects for reducing the processing cost of lithium ion batteries. *J. Power Sources* **275**, 234–242 (2015).
30. Zimmerman, A. H., Quinzio, M. V. & Monica, S. Adaptive charging method for lithium-ion battery cells. US Patent US6204634B1 (2001).
31. Park, S., Kato, D., Gima, Z., Klein, R. & Moura, S. Optimal experimental design for parameterization of an electrochemical lithium-ion battery model. *J. Electrochem. Soc.* **165**, A1309–A1323 (2018).
32. Smith, J. S., Nebgen, B., Lubbers, N., Isayev, O. & Roitberg, A. E. Less is more: sampling chemical space with active learning. *J. Chem. Phys.* **148**, 241733 (2018).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Methods

Experimental

Commercial high-power lithium iron phosphate (LFP)/graphite A123 APR18650M1A cylindrical cells were used in this work (packing date 2015-09-26, lot number EL1508007-R). These cells have a nominal capacity of 1.1 A h and a nominal voltage of 3.3 V. All currents are defined in units of C rate; here, 1C is 1.1 A, or the current required to fully (dis)charge the nominal capacity (1.1 A h) in 1 h. The manufacturer's recommended fast-charging protocol is 3.6C (3.96 A) CC-CV. The rate capability of these cells is shown in Extended Data Fig. 2. The graphite and LFP electrodes are 40 μm thick and 80 μm thick, respectively, as quantified via X-ray tomography (Zeiss Xradia 520 Versa).

The cells were cycled with various charging protocols but identically discharged. Cells were charged with one of 224 candidate six-step, ten-minute charging protocols from 0% to 80% SOC, as detailed below. After a five-second rest, all cells then charged from 80% to 100% SOC with a 1C CC-CV charging step to 3.6 V and a current cutoff of C/20. After another five-second rest, all cells subsequently discharged with a CC-CV discharge at 4C to 2.0 V and a current cutoff of C/20. The cells rested for another five seconds before the subsequent charging step started. The lower and upper cutoff voltages were 2.0 V and 3.6 V, respectively, as recommended by the manufacturer. In this work, cycle life is defined as the number of cycles until the discharge capacity falls below 80% of the nominal capacity.

All cells were tested in cylindrical fixtures with 4-point contacts on a 48-channel Arbin Laboratory Battery Testing battery cyclor placed in an environmental chamber (Amerex Instruments) at 30 °C. The cyclor calibration was validated before the state of the experiment.

In the closed-loop experiment, four experiments did not reach 100 cycles owing to contact issues either at the start or partially through the experiment. These experiments were run on channels 17 and 27 in round 1 (oed_0) and channels 4 and 5 in round 2 (oed_1). Additionally, in each round, one protocol per round that should have been selected (that is, with a top-48 upper bound) was not selected and replaced with the protocol with the 49th-highest upper bound owing to a processing error (Extended Data Fig. 4), but this error is not expected to have a large effect. Additional experimental issues are documented in the notes of the data repository.

Charging protocol and parameter space design

Cells were charged with one of 224 different four-step charging protocols. Each of the first four steps is a single constant-current step applied over a 20% SOC range; thus, the 224 charging protocols represent different combinations of current steps within the 0% to 80% SOC range. We can define the charging time from 0% to 80% SOC by:

$$t_{0-80\%} = \frac{0.2}{CC1} + \frac{0.2}{CC2} + \frac{0.2}{CC3} + \frac{0.2}{CC4}$$

In all protocols considered here, we constrain $t_{0-80\%}$ to be 10 min. We now write CC4 as a function of the first three charging steps, as:

$$CC4 = \frac{0.2}{\frac{10}{60} - \left(\frac{0.2}{CC1} + \frac{0.2}{CC2} + \frac{0.2}{CC3} \right)}$$

Thus, each protocol can be uniquely defined by CC1, CC2 and CC3.

Each independent parameter can take on one of the following discrete values: 3.6C, 4.0C, 4.4C, 4.8C, 5.2C and 5.6C. Furthermore, CC1 can take on values of 6.0C, 7.0C and 8.0C, and CC2 can take on values of 6.0C and 7.0C. CC4 is not allowed to exceed 4.81C. The maximum allowable current for each parameter decreases with increasing SOC to avoid reaching the upper cutoff voltage of 3.6 V. With these constraints, a total of 224 charging protocols are permitted.

For a consistent protocol nomenclature, we define each fast-charging protocol as CC1-CC2-CC3-CC4. For example, the charging protocol with the highest CLO-estimated mean cycle life is written 4.8C-5.2C-5.2C-4.160C.

Early outcome predictor

The early outcome predictor for cycle life is similar to that presented in Severson et al.⁵. This linear model predicts the final \log_{10} cycle life (number of cycles to reach 80% of nominal capacity, or 0.88 A h) using features from the first 100 cycles. The training set is identical to the one used in Severson et al.⁵ and consists of 41 batteries. The linear model takes the form:

$$\hat{y}_i = \hat{\mathbf{w}}^T \mathbf{x}_i$$

Here \hat{y}_i is the predicted cycle life for battery i , \mathbf{x}_i is a p -dimensional feature vector for battery i and $\hat{\mathbf{w}}$ is a p -dimensional model coefficient vector. Features are z-scored (mean-subtracted and normalized by the standard deviation) to the training set before model evaluation.

Regularization, or simultaneous feature selection and model fitting, was performed using the elastic net²⁷. Regularization penalizes overly complex fits to improve both generalizability and interpretability. Specifically, the coefficient vector $\hat{\mathbf{w}}$ is found via the following expression:

$$\hat{\mathbf{w}} = \underset{\mathbf{w}}{\text{argmin}} \left[\|\mathbf{y} - \mathbf{X}\mathbf{w}\|_2^2 + \lambda \left(\frac{1-\alpha}{2} \|\mathbf{w}\|_2^2 + \alpha \|\mathbf{w}\|_1 \right) \right]$$

Here λ and α are hyperparameters; λ is a non-negative scalar and α is a scalar between 0 and 1. The first term minimizes the squared loss, and the second term performs both continuous shrinkage and automatic feature selection. During model development, we apply fourfold cross-validation and Monte Carlo sampling with the training set to optimize the values of the hyperparameters λ and α .

As in Severson et al.⁵, the available features were based on the difference between discharge voltage curves of cycles 100 and 10, or trends in the discharge capacity. The five selected features, their corresponding weights and the z-scored values are presented in Supplementary Table 1. The training (cross-validated) error was 80.4 cycles (10.2%); the test error on a test set from Severson et al.⁵ was 122 cycles (12.6%).

The early predictor automatically flags predictions as anomalous if the 95% prediction interval exceeds 2,000 cycles. The two-tailed 95% prediction interval is computed by:

$$95\%PI = 2t_{(\alpha/2, n-p)} \times RMSE \sqrt{1 + \mathbf{x}_i^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{x}_i}$$

where t is the Student's t value, α is the significance level (0.05 for 95% confidence), n is the number of samples, p is the number of features, RMSE is the root-mean-square error of the training set (in units of cycles), \mathbf{x}_i is the vector of selected features for battery i and \mathbf{X} is the matrix of selected features for all observations in the training set.

In the closed-loop experiment, three tests returned predictions with a prediction interval outside of the threshold; these anomalous predictions were excluded. These tests were run on channel 27 in round 1 (oed_0), channel 12 in round 3 (oed_2) and channel 6 in round 4 (oed_3). Furthermore, in the validation experiment, one test returned a prediction with a prediction interval outside of the threshold (channel 12; 3.6C-6.0C-5.6C-4.755C), although the final cycle life was reasonable.

We note that the predictions from this model exhibited systematic bias for the cells in the validation experiments, which we attribute to the increased calendar ageing of these cells relative to the training set (Supplementary Table 2 and Supplementary Discussion 4).

Bayesian optimization algorithm

To perform optimal experimental design, we consider the setting of best-arm identification using multi-armed bandits. Here each arm is a charging protocol and the goal is to identify the best arm, or equivalently the charging protocol with the highest expected cycle life. Many variants of the problem have been studied in prior work^{33–35}; our algorithm builds on the approaches of Hoffman et al.¹⁰ and Grover et al.¹¹. We consider further modifications in Supplementary Discussion 2.

In particular, we assume a Bayesian regression setting, where there exists an unknown set of parameters ($\theta \in \mathbb{R}^d$) that relate a charging protocol x to its cycle life (a scalar) via a Gaussian likelihood function. Here, x denotes the CC1, CC2, CC3 configurations of a charging protocol, which is projected onto a d -dimensional feature vector $\phi(x)$. We set $d = 224$, and the feature representations $\phi(x)$ are obtained by approximating a radial-basis function kernel, $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|_2^2)$, using Nystroem's method. Here, x_i and x_j are the CC1, CC2 and CC3 configurations for two arbitrary charging protocols and the inverse of the kernel bandwidth, $\gamma > 0$ is treated as a hyperparameter.

The Gaussian likelihood function relates a charging protocol to its cycle life distribution. For a protocol x , the mean of this likelihood function is given as $\theta^T \phi(x)$. The variance of this likelihood function is the sum of two uncertainty terms, both of which we assume to be homoskedastic (that is, uniform across all protocols). The first term is the empirical variance averaged across the repeated runs of individual protocols present in the training dataset (same as that used for training the early predictor). This accounts for variability due to exogenous factors such as manufacturing. Second, since we do not wait for an experiment to complete, the likelihood variance additionally needs to accommodate an additional uncertainty term due to the early outcome prediction component of the pipeline. We do so by computing the residual variance of the early predictions on the held-out portion of the dataset and set the aforementioned uncertainty term to be the maximum of the residual variances. We assume that the two sources of uncertainty are independent, and hence the overall variance of the likelihood distribution is given by the sum of the squares of both variance terms described above.

To perform inference over the unknown parameters θ and subsequent predictions of cycle lives, we employ a Gaussian process. In a Gaussian process, the prior over θ is assumed to be isotropic Gaussian; such a prior is conjugate to the Gaussian likelihood, and as a consequence the Gaussian posterior can be obtained in closed-form via the Bayes rule. This posterior is used to define a Gaussian predictive distribution over the cycle life for any given charging protocol with mean μ and variance σ^2 .

Finally, to select a charging protocol, we optimize an acquisition function based on upper confidence bounds. The acquisition function selects protocols where the noisy predictive distribution over cycle life has high mean μ (to encourage exploitation) and high variance σ^2 (to encourage exploration). The mean and upper and lower confidence bounds for any arm i is given by $\mu_{k,i} \pm \beta_k \sigma_{k,i}$ at round k , such that the relative weighting of the two terms is controlled by the exploration tradeoff hyperparameter, $\beta > 0$. The exploration tradeoff hyperparameter at round k , β_k , is decayed multiplicatively at every round of the closed loop by another hyperparameter, $\varepsilon \in (0,1]$, as given by $\beta_k = \beta_0 \varepsilon^k$.

BO hyperparameter optimization

The BO algorithm relies on eight hyperparameters, each of which can be categorized as either a resource hyperparameter, a parameter space hyperparameter or an algorithm hyperparameter. We note that the BO algorithm runs in the fixed-budget setting; here, the budget refers to the number of iterations of the closed loop we run, excluding validation experiments. We describe each category of hyperparameters below; the values of each hyperparameter are tabulated in Supplementary Table 5.

Resource hyperparameters are specified by the available testing resources. The 'batch size' represents the number of parallel tests. We set a batch size of 48 given our 48-channel battery cycler. The 'budget' represents the number of batches tested during CLO. The budget excludes batches used to develop the early predictor and validation batches. The budget is typically constrained by either the available testing time or the number of cells. In this case, we set a budget of 4, yielding a cell budget of 192 cells and a time budget of 16 days (4 days per batch of 48 cells tested for 100 cycles).

Parameter space hyperparameters are specified by the optimization problem. Here, we use the same data available from the training set of the early predictor to estimate these parameters, despite a different charging protocol structure. The 'standardization mean' represents the estimated mean cycle life across all protocols. The 'standardization standard deviation' represents the estimated standard deviation of cycle life across all protocols; in other words, this parameter represents the range of cycle lives in the parameter space. The 'likelihood standard deviation' represents the estimated standard deviation of a single protocol tested multiple times, which is a measure of the sampling error; this sampling error includes both the intrinsic variability and the prediction error.

Algorithm hyperparameters control the performance of the Bayesian optimization algorithm. γ is the kernel bandwidth, which controls the interaction strength between neighbouring protocols in the parameter space. High γ favours under-smoothing of the parameter space, that is, the protocols have weak relationships with their neighbours. β_0 represents the initial value of β , the exploration tradeoff hyperparameter; β controls the balance of exploration versus exploitation. High β_0 favours exploration over exploitation. ε represents the decay constant of beta per round; as the experiment progresses, ε shifts towards stronger exploitation (given by $\beta_k = \beta_0 \varepsilon^k$, where β_k represents the exploration constant at round k , 0-indexed). High ε favours a rapid transition from exploration to exploitation.

The algorithm hyperparameters were estimated by creating a physics-based simulator based on the range of cycle lives obtained in the preliminary batch, testing all hyperparameter combinations on the simulator, and selecting the hyperparameter combination with the best performance (that is, that which most consistently obtains the true cycle life). These results are visualized in Extended Data Fig. 9; we note that the performance of BO is relatively insensitive to the selected combination of algorithm hyperparameters, meaning sufficiently high performance can be achieved even with suboptimal algorithm hyperparameters. Other approaches, such as using the early-predictor training dataset, are also possible for optimization of the algorithm hyperparameters (Supplementary Discussion 1).

Physics-based simulator

We used a physics-based simulator for hyperparameter optimization; this simulator allows us to estimate the shape and range of cycle lives in the parameter space, although the simulator is not designed to be an accurate representation of battery degradation during fast charging. This finite element simulator was originally designed to estimate the heat generation during charging in an 18650 cylindrical battery by approximating the battery as a long cylinder, which simplifies to a one-dimensional radial heat transfer problem. The equations and thermal properties were sourced from Drake et al.³⁶ and Çengel and Boles³⁷. The output from these simulations is a matrix of temperature as a function of both radial position and time. We use total solid-electrolyte interphase (SEI) growth as a proxy for degradation. First, we estimate the temperature dependence of SEI growth from the C/10 series of figure 7 from Smith et al.³⁸ (Supplementary Table 6). Simultaneously, we compute the expected temperature profiles in the battery as a function of charging protocol with respect to time and position. We then approximate the kinetics of SEI growth with an Arrhenius equation, such that SEI growth increases with increasing temperature. SEI growth

(in arbitrary units) is calculated for each temperature element in the position-time array via:

$$D = \sum_r \sum_t \exp\left(-\frac{E_a}{k_B T}\right)$$

where D is the degradation parameter, E_a is the effective activation energy for SEI growth (Supplementary Table 6) and k_B is Boltzmann's constant. The cycle life is then calculated from the degradation parameter using the range of expected cycle lives (as estimated from the early-predictor training dataset):

$$\text{Cycle life} = 500 + C/D$$

where C is a constant (5×10^{-11}) that scales D to reasonable values of cycle life.

Validation experiments

After the closed-loop experiment completed, we selected nine protocols to test to failure (five batteries per charging protocol). This experiment allowed us to (1) evaluate the performance of the closed loop by comparing the CLO-estimated mean cycle lives to the mean cycle life of multiple batteries tested to failure for multiple protocols, (2) compare the protocols with the highest CLO-estimated mean cycle lives to conventional fast-charging protocol design principles from the battery literature, and (3) generate a small dataset with which we can evaluate the performance of the closed loop relative to baseline optimization approaches.

The selected protocols are displayed in Extended Data Fig. 6 and Extended Data Table 1. Of our nine fast-charging protocols, three were the top three CLO-estimated protocols; four were based on approximations of multi-step fast-charging protocols in the battery literature (see Extended Data Table 1); and two were selected to obtain a representative sampling from the distribution of CLO-estimated cycle lives. The four protocols based on approximations of multi-step fast-charging protocols in the battery literature were obtained by determining the current ratios between various steps and translating those ratios to our ten-minute fast-charging space. The voltage limits were consistent with our charging protocols, that is, 2.0 V and 3.6 V.

Five batteries per charging protocol were tested to obtain a reasonable estimate of the true cycle lives. In this experiment, one test returned a prediction with a prediction interval outside of the threshold (channel 12; 3.6C-6.0C-5.6C-4.755C) and was excluded. A comparison of the three different methods for cycle life results (CLO, early predictions from validation, and final measurements from validation) are presented in Extended Data Fig. 7.

Validation ablation study

For the ablation study using the charging protocols and data from the validation experiments, we systematically compared the full closed-loop system against three other ablation baselines which use (1) only early prediction (no BO exploration-exploitation, purely random exploration), (2) only BO exploration-exploitation (no early prediction), (3) purely random exploration without any early prediction. As highlighted earlier, since the final cycle lives for the protocols in the validation study have a noticeable bias that can be explained by calendar ageing (Supplementary Discussion 4), we perform a simple additive bias correction for each of the final cycle lives beforehand to suppress any undesirable influence of this bias in interpreting the results.

We run the four ablation baselines for a varying number of sequential rounds. Since our validation space is relatively small (nine charging protocols, five batteries tested per protocol in our validation dataset), we run only one battery per round (that is, we assume a one-channel battery cyler). The baselines that use BO exploration-exploitation additionally require hyperparameters to be specified before beginning

the experiment, as described in the Methods section 'BO hyperparameter optimization'. The best hyperparameters are chosen separately for each round based on the performance obtained on the physics-based simulator, averaged over 100 random seeds.

When an ablation baseline queries for the cycle life of a given charging protocol, the returned value corresponds to one of the five runs in our validation dataset, chosen via random sampling with replacement (that is, bootstrapped). The experimental time cost of this query is equal to 100 cycles for ablation baselines that use early prediction and equals the full cycle life otherwise. Finally, to account for the randomness at the beginning of the experiment (that is, round 0 when every ablation baseline randomly selects a protocol), we report the performance of each ablation baseline averaged over a sequence of 2,000 randomly initialized experiments. To specify the y-axis of Fig. 4e, we assume that each full cycle (charging, discharging, resting) requires one hour of experimental testing.

Overpotential analysis

To determine the dependence of overpotential on current and SOC during charging (Extended Data Fig. 2e-f), we perform a pseudo-galvanostatic intermittent titration technique experiment on two minimally cycled batteries and two degraded batteries (80% of nominal capacity remaining). We probe currents ranging from 3.6C to 8C at 20%, 40%, 60% and 80% SOC, mirroring the current and SOC values used in charging protocol design. In this experiment, we start at an initial SOC 20% lower than the target, for example, we start at 0% SOC to probe 20% SOC. We then charge at a given current rate, for example, 3.6C, until we reach 20% SOC. The cell rests for 1 h, and then the cell discharges at 1C back to 0% SOC. We repeat this sequence for all current values, after which we charge the cell at 1C to the next initial SOC, for example, 20% SOC to probe 40% SOC, and repeat for each SOC of interest.

To compute the overpotential, we compare the voltage at the start and end of the 1-h rest periods. Nearly all of the potential drop occurs immediately (<100 ms) after the start of the rest period. Given the linear trends observed (implying ohmic-limited rate capability), we then perform a linear fit on each overpotential-current series. In these fits, the slope represents the ohmic resistance.

Data availability

The datasets used in this study are available at <https://data.matr.io/1>.

Code availability

The CLO code, data and figures associated with this manuscript are available at <https://github.com/chueh-ermon/battery-fast-charging-optimization>. The data processing and early-prediction code are available at <https://github.com/chueh-ermon/BMS-autoanalysis>. The charging protocol generation code (automated creation of battery cyler tests) is available at <https://github.com/chueh-ermon/automate-Arbin-schedule-file-creation>.

33. Shahriari, B., Swersky, K., Wang, Z., Adams, R. P. & de Freitas, N. Taking the human out of the loop: a review of Bayesian optimization. *Proc. IEEE* **104**, 148–175 (2016).
34. Audibert, J.-Y., Bubeck, S. & Munos, R. Best arm identification in multi-armed bandits. In *Proc. 23rd Conf. on Learning Theory (COLT)* 41–53 (2010); <http://certis.enpc.fr/~audibert/Mes%20articles/COLT10.pdf>.
35. Srinivas, N., Krause, A., Kakade, S. M. & Seeger, M. W. Information-theoretic regret bounds for Gaussian process optimization in the bandit setting. *IEEE Trans. Inf. Theory* **58**, 3250–3265 (2012).
36. Drake, S. J. et al. Measurement of anisotropic thermophysical properties of cylindrical Li-ion cells. *J. Power Sources* **252**, 298–304 (2014).
37. Çengel, Y. A. & Boles, M. A. *Thermodynamics: An Engineering Approach* (McGraw-Hill Education, 2015).
38. Smith, A. J., Burns, J. C., Zhao, X., Xiong, D. & Dahn, J. R. A high precision coulometry study of the SEI growth in Li/graphite cells. *J. Electrochem. Soc.* **158**, A447–A452 (2011).
39. Zhang, S. S. The effect of the charging protocol on the cycle life of a Li-ion battery. *J. Power Sources* **161**, 1385–1391 (2006).

40. Kim, J. M. et al. Battery charging method and battery pack using the same. US Patent Application US20160226270A1 (2016).
41. Lee, M.-S., Song, S.-B., Jung, J.-S. & Golovanov, D. Battery charging method and battery pack using the same. US Patent US9917458B2 (2018).
42. Notten, P. H. L., Op het Veld, J. H. G. & van Beek, J. R. G. Boostcharging Li-ion batteries: a challenging new charging concept. *J. Power Sources* **145**, 89–94 (2005).
43. Paryani, A. Low temperature charging of Li-ion cells. US Patent US8552693B2 (2013).
44. Mehta, V. H. & Straubel, J. B. Fast charging with negative ramped current profile. US Patent US8643342B2 (2014).

Acknowledgements This work was supported by the Toyota Research Institute through the Accelerated Materials Design and Discovery programme. P.M.A. was supported by the Thomas V. Jones Stanford Graduate Fellowship and the National Science Foundation Graduate Research Fellowship under grant number DGE-114747. A.G. was supported by a Microsoft Research PhD Fellowship and a Stanford Data Science Scholarship. N.P. was supported by the SAIC Innovation Center through the Stanford Energy 3.0 industry affiliates programme. S.J.H. was supported by the Assistant Secretary for Energy Efficiency, Vehicle Technologies Office of the US Department of Energy under the Advanced Battery Materials Research Program. X-ray tomography was performed at the Stanford Nano Shared Facilities, supported by the National Science Foundation under award ECCS-1542152. We thank A. Anapolsky, L. Attia, C. Cundy, J.

Hirshman, S. Jorgensen, G. McConohy, J. Song, R. Smith, B. Storey and H. Thaman for discussions.

Author contributions P.M.A., N.J., Y.-H.L., M.H.C., N.P. and W.C.C. conceived and conducted the experiments. A.G., T.M.M., B.C. and S.E. developed the Bayesian optimization algorithm and incorporated early outcome predictions into the closed loop. P.M.A. and K.A.S. performed the early-prediction modelling. P.M.A., Z.Y., P.K.H. and M.A. performed data management. P.M.A., A.G., N.J., S.J.H., S.E. and W.C.C. interpreted the results. All authors edited and reviewed the manuscript. R.D.B., S.E. and W.C.C. supervised the work.

Competing interests S.E., W.C.C., A.G., T.M.M., N.P. and P.M.A. have filed a patent application related to this work: US Patent Application No. 16/161,790 (16 October 2018).

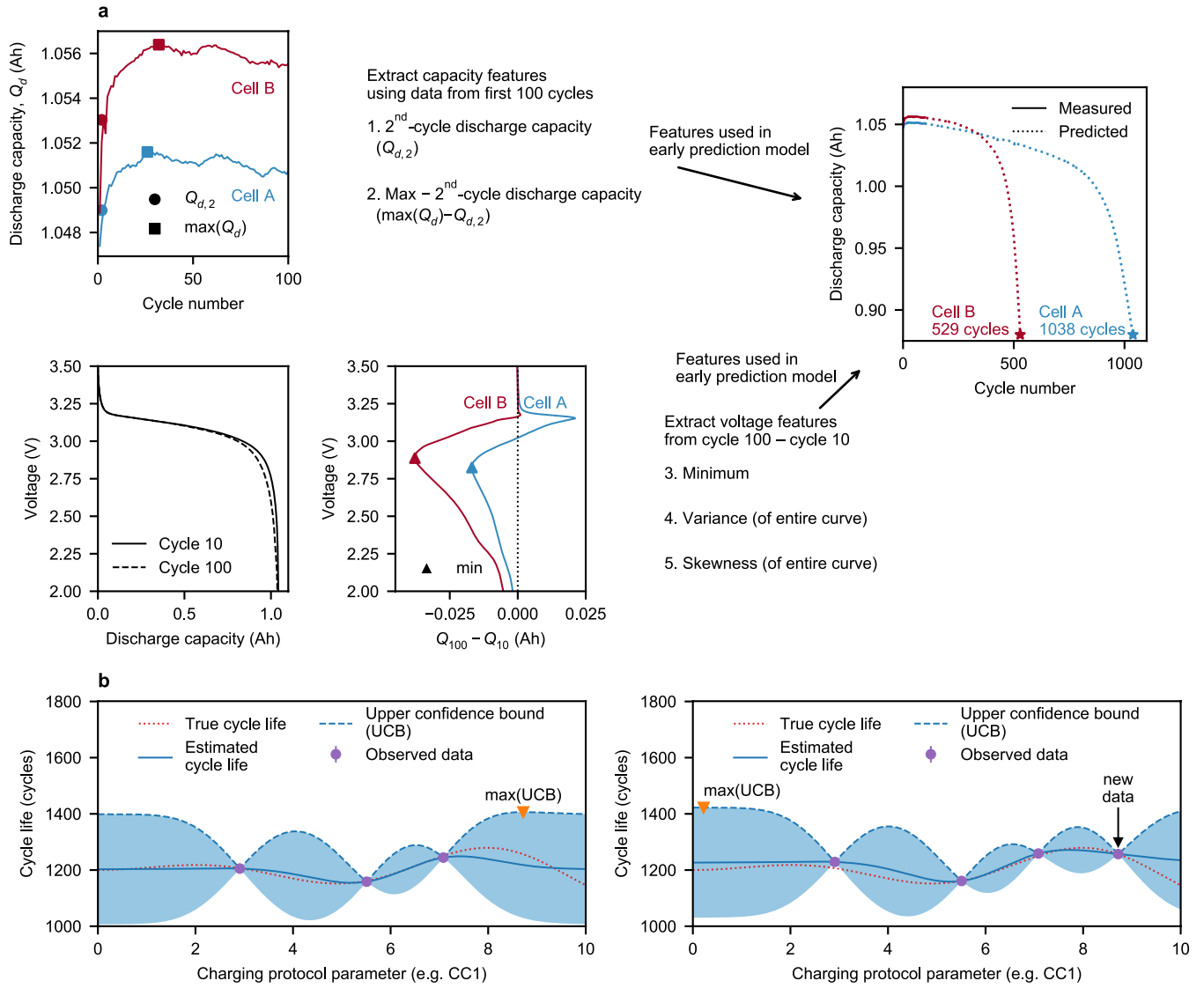
Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-1994-5>.

Correspondence and requests for materials should be addressed to R.D.B., S.E. or W.C.C.

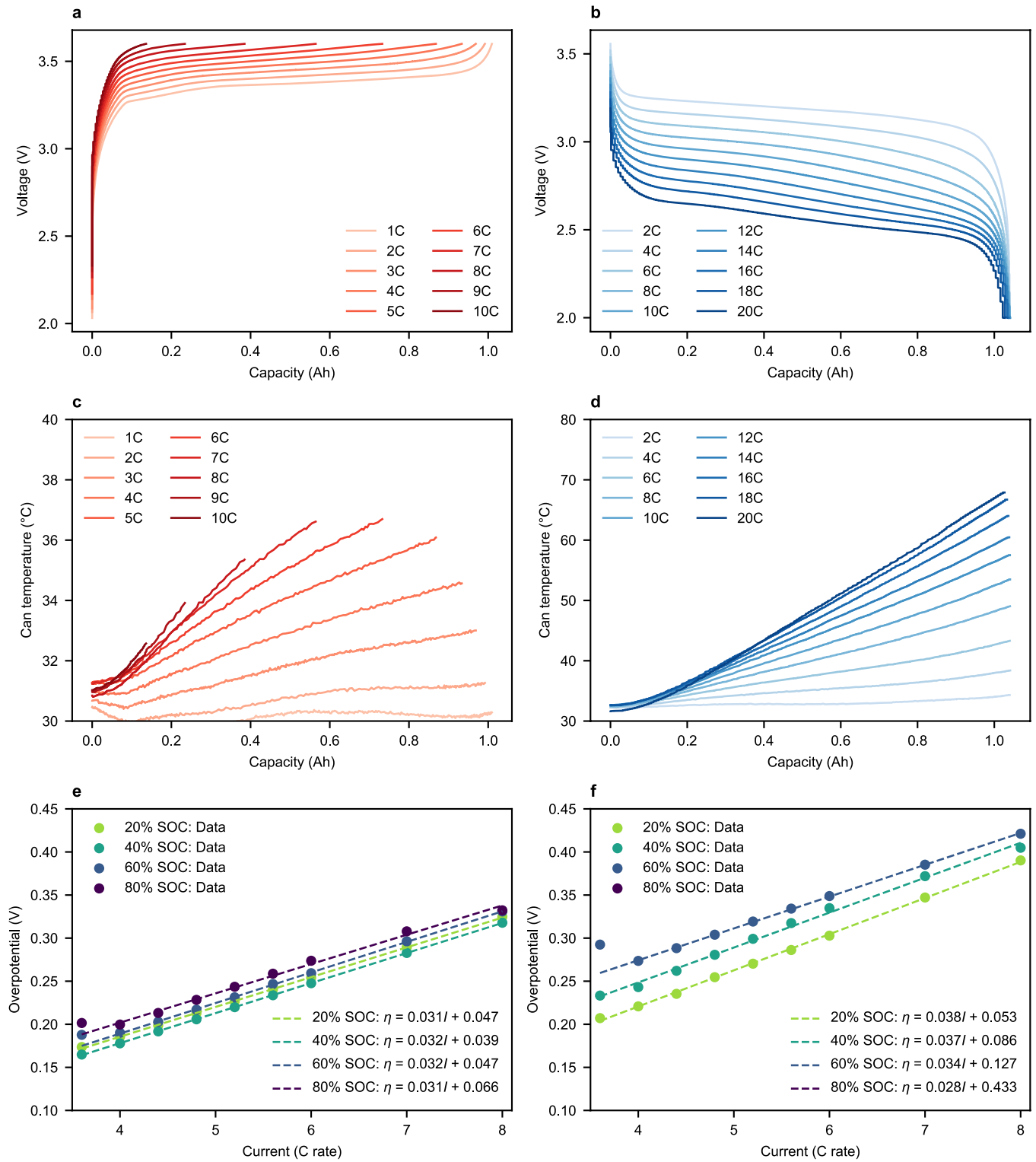
Peer review information *Nature* thanks Marius Bauer, Matthias Seeger and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



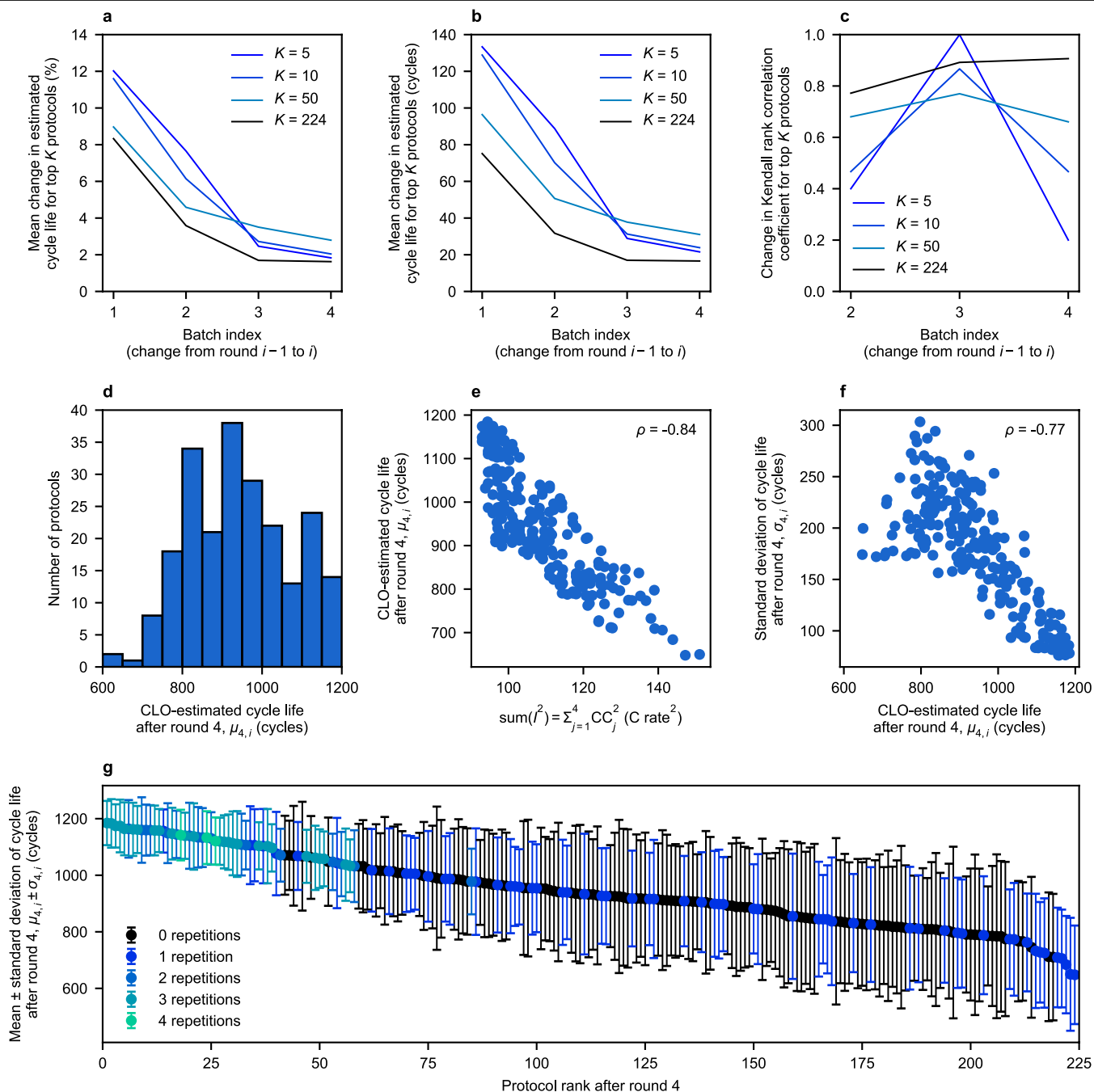
Extended Data Fig. 1 | Illustrations of early outcome predictor and BO components of CLO. a, Illustration of early outcome prediction for two cells (A and B) using data from only the first 100 cycles. Two discharge capacity features are generated: the second-cycle discharge capacity, $Q_{d,2}$, and the difference between the maximum and second-cycle discharge capacities, $\max(Q_d) - Q_{d,2}$. Three voltage features are generated: the logarithm of the minimum, variance and the skewness of the difference in voltage curves between cycles 100 and 10. These five features are combined in a linear model to predict the final cycle life, or the number of cycles until the capacity falls below 0.88 A h. The weights and scalings of each feature are determined by training the model on a training set using the elastic net; the weights and scaling values are presented in Supplementary Table 1. See Severson et al.⁷ and

Methods for additional details. **b,** Illustration of the BO principle. The desired output, cycle life, has a true functional dependence on charging protocol parameters (such as CC1). Here, we show a one-dimensional model (that is, just dependent on one parameter, CC1) for simplicity. By performing Gaussian process regression on the available data, we develop a probabilistic estimate of the true function; our goal is for the estimate to match the true function. The next data point selected is that which maximizes the upper confidence bound (UCB), which is selected by either high uncertainty (exploration) or high predicted value (exploitation). Once this point is selected (right panel), the next point selected is, again, that which maximizes the upper confidence bound.



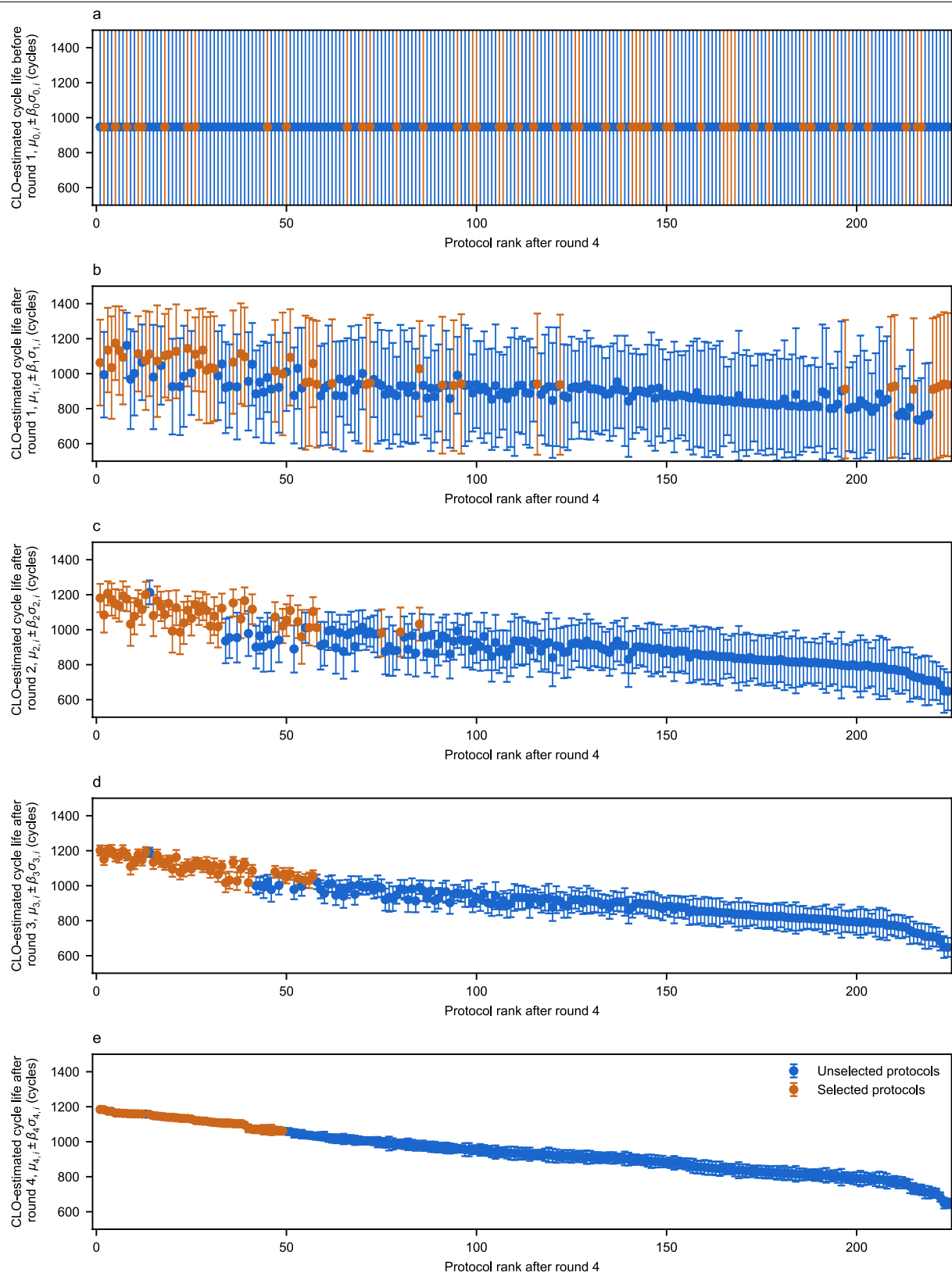
Extended Data Fig. 2 | Cell characterization. **a, b**, Voltage versus capacity during rate testing of A123 18650M1A cylindrical cells under charge (**a**) and discharge (**b**). The (dis)charge step not under investigation is cycled at 1C to isolate the rate of each step; for example, the charge rate test is performed with 1-C discharge steps. We note that the discharge rate capability is much higher than that of charge. **c, d**, Battery surface temperature ('can temperature') versus capacity during rate testing under charge (**c**) and discharge (**d**). The can temperature is measured via a type T thermocouple secured with thermal

epoxy. **e, f**, Overpotential as a function of SOC and C rate (see Methods section 'Overpotential analysis' for details of the measurement) for a minimally cycled cell (**e**) and an aged cell at 80% of nominal capacity (**f**). The trend lines are linear fits of the overpotential as a function of current at fixed SOC (excluding outliers). We note that both of the relationships are linear (indicating that the rate capability is ohmically limited) and that the SOC dependence is weak, particularly for the minimally cycled cell. The initial internal resistance, averaged over two cells and all four SOC, is 33 mΩ.



Extended Data Fig. 3 | Additional optimization results. **a, b**, Mean of the absolute difference in CLO-estimated cycle lives with increasing rounds, expressed as both percentage change (**a**) and absolute change (**b**). These changes are relatively small beyond round 2, suggesting that the closed loop can perform well with even smaller time or battery budgets. **c**, Change in Kendall rank correlation coefficient with increasing rounds. From round 3 to round 4, the ranking of the top protocols shifts, but the cycle lives of these top protocols are similar. **d**, Distribution of CLO-estimated mean cycle lives after round 4. The mean and standard deviation are 943 cycles and 126 cycles, respectively. **e**, Correlation between CLO-estimated mean cycle lives and the sum of squared currents, a simplified measure of heat generation ($P = I^2 R$). This relationship suggests that minimizing heat generation, as opposed to avoiding

lithium plating, may be the operative optimization strategy for these cells under these conditions. **f**, Standard deviation ($\sigma_{4,i}$) versus mean ($\mu_{4,i}$) of the BO predictive distribution over cycle life after round 4. The standard deviation quantifies the uncertainty in the cycle life estimates and is generally low for protocols estimated to have high mean cycle life, since these protocols are probed more frequently. We start with a relatively wide, flat prior (standard deviation 164) and therefore the uncertainty intervals after four rounds are also wide. **g**, Mean \pm standard deviation of the predictive distribution over cycle life after round 4 ($\mu_{4,i} \pm \sigma_{4,i}$) for all charging protocols, sorted by their rank after round 4. The legend indicates the number of repetitions for each protocol (excluding failed batteries).

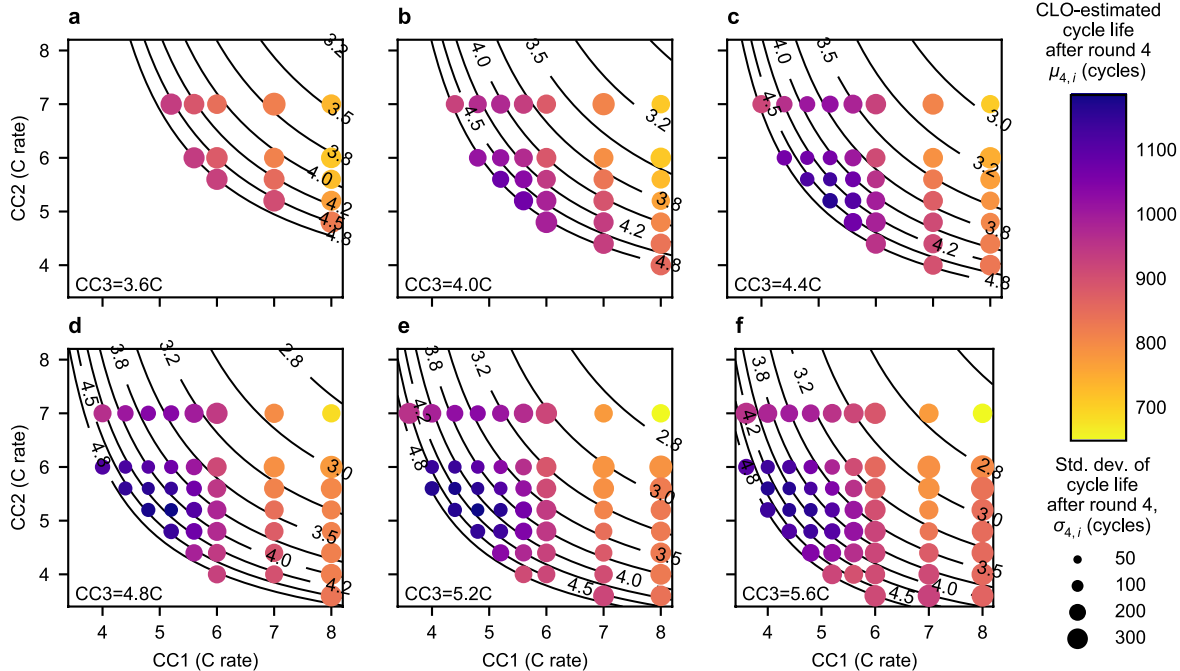


Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Means and upper/lower confidence bounds

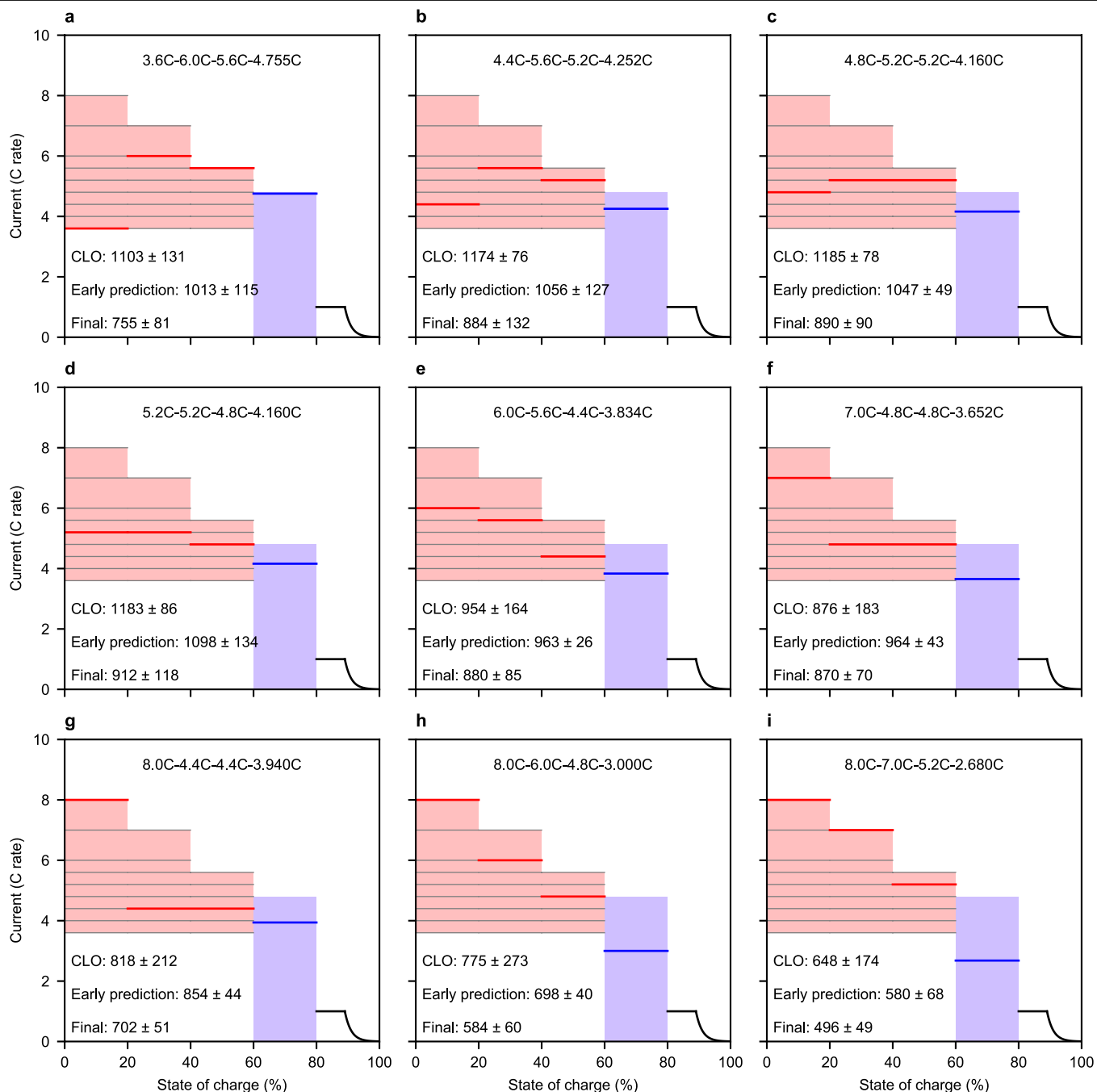
($\mu_{k,i} \pm \beta_k \sigma_{k,i}$) on cycle life per round k . Protocol indices on the x -axis are sorted by rank after round 4. The weighted interval around the estimated mean, $\beta_k \sigma_{k,i} = (\beta_0 \varepsilon^k) \sigma_{k,i}$, weights the protocol-specific standard deviation at round k , $\sigma_{k,i}$ (estimated by the Gaussian process model) with the exploration tradeoff hyperparameter at round k , β_k . The upper and lower confidence bounds are plotted for all charging protocols before round 1 (**a**) and after rounds 1 (**b**), 2 (**c**), 3 (**d**) and 4 (**e**). The predictive distributions for all charging protocols have identical means and standard deviations before the first round of testing. Because the standard deviations are weighted by $\beta_k = \beta_0 \varepsilon^k$ and $\varepsilon = 0.5$, the

weighted confidence bounds rapidly decrease with increasing round number, favouring exploitation (examination of protocols with high means). The BO algorithm recommends the 48 protocols with the highest upper bounds (red points); the upper bounds are high either due to high uncertainty (exploration) or high means (exploitation). The algorithm rapidly shifts from exploration to exploitation as ε_k rapidly shrinks the upper bounds with increasing round index. We note that one protocol per round that should have been selected (that is, with a top-48 upper bound) was not selected owing to a processing error; instead, the protocol with the 49th-highest upper bound was selected.



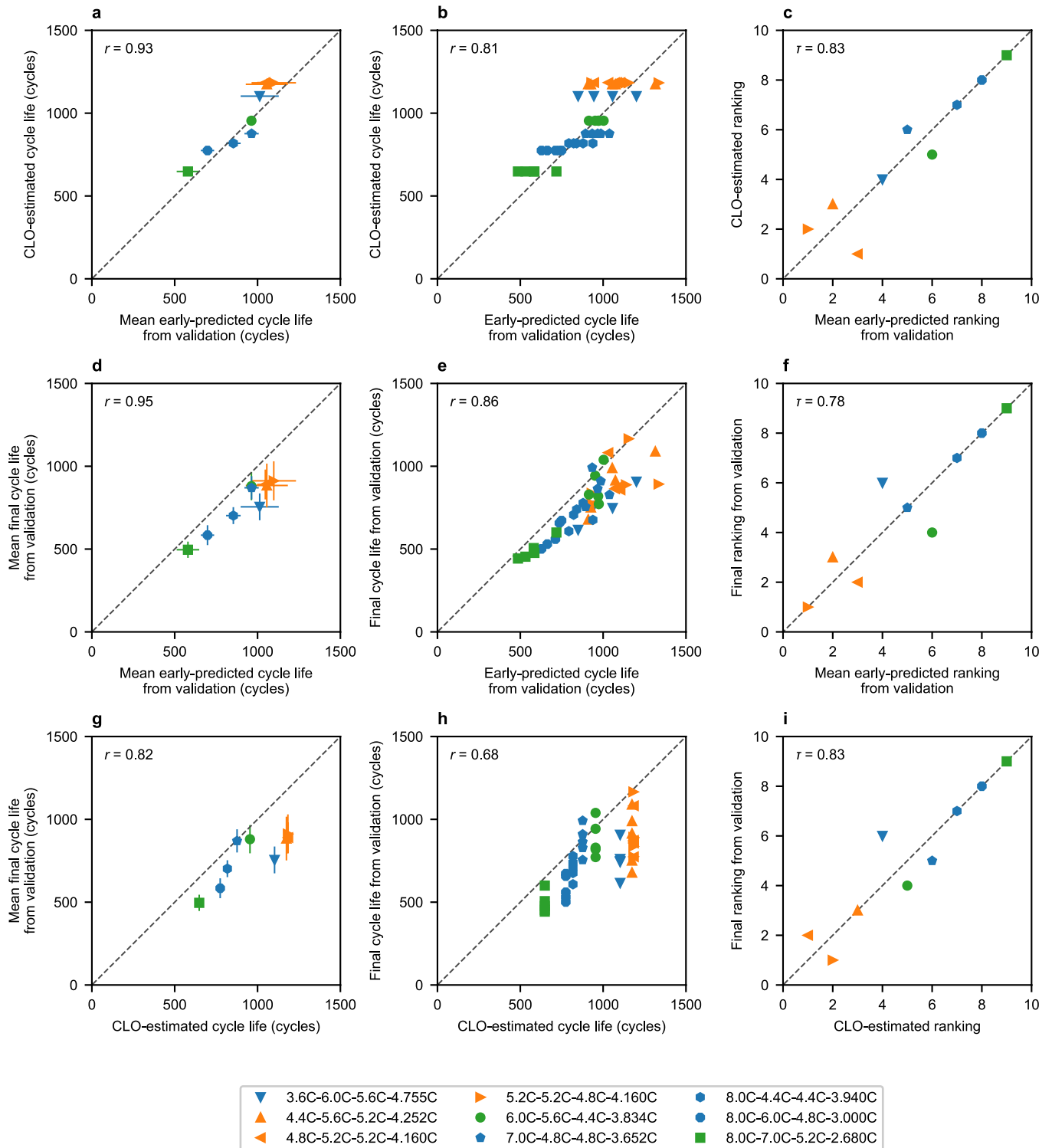
Extended Data Fig. 5 | Mean and standard deviation of the CLO-estimated predicted distribution over cycle lives after round 4. In this two-dimensional representation, mean estimated cycle life (colour scale) and standard deviation of cycle life (marker size) after round 4 are presented as a function of CC1, CC2 and CC3 (the x axis, y axis and panels a–f, respectively). Panels a–f represent

CC3=3.6C, 4.0C, 4.4C, 4.8C, 5.2C, 5.6C and 6.0C, respectively. CC4 is represented by the contour lines. Note that the protocols with the highest cycle lives generally have the smallest standard deviations, since these protocols have been tested repeatedly.



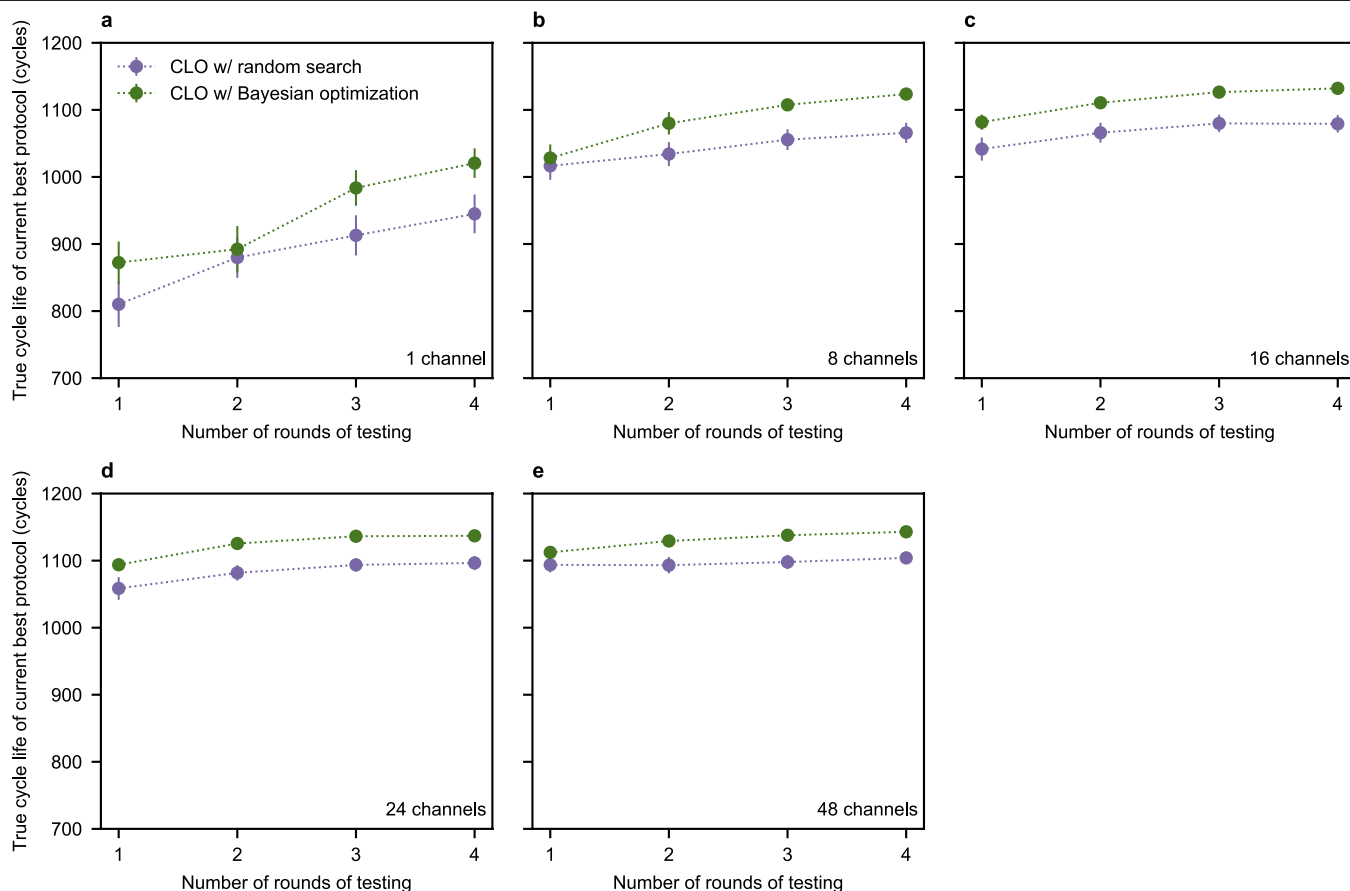
Extended Data Fig. 6 | Selected protocols for validation. The three protocols with the highest CLO-estimated mean cycle lives are shown in panels **b**, **c** and **d**. The protocols shown in panels **a**, **f**, **g** and **h** are approximations of previously proposed battery fast-charging protocols (Extended Data Table 1). The remaining two protocols, shown in panels **e** and **i**, were selected to obtain a representative sampling from the entire distribution of CLO-estimated cycle lives. The annotations on each panel represent the cycle lives of each protocol

as estimated by CLO ('CLO'), early outcome prediction from validation ('Early prediction'), and the final cycle lives from validation ('Final'). In the annotations, the errors represent the CLO-estimated standard deviation after round 4 ($\sigma_{k,4}$) for the CLO-estimated cycle lives and the 95% confidence intervals for the early-predicted and final cycle lives from validation ($n=5$; $n=4$ for the early predictions of 3.6C-6.0C-5.6C-4.755C) (**a**).



Extended Data Fig. 7 | Validation ablation analysis. We perform pairwise comparisons of the cycle lives of the nine validation protocols, as estimated from three sources: closed-loop estimates after four rounds, early predictions from the validation experiment and final cycle lives from the validation experiment. Panels **a–c** compare closed-loop estimates to early predictions from validation, panels **d–f** compare final cycle lives from validation to early predictions from validation, and panels **g–i** compare final cycle lives from validation to closed-loop estimates. The first column (**a**, **d** and **g**) compares cycle lives averaged on a protocol basis; the second column (**b**, **e** and **h**)

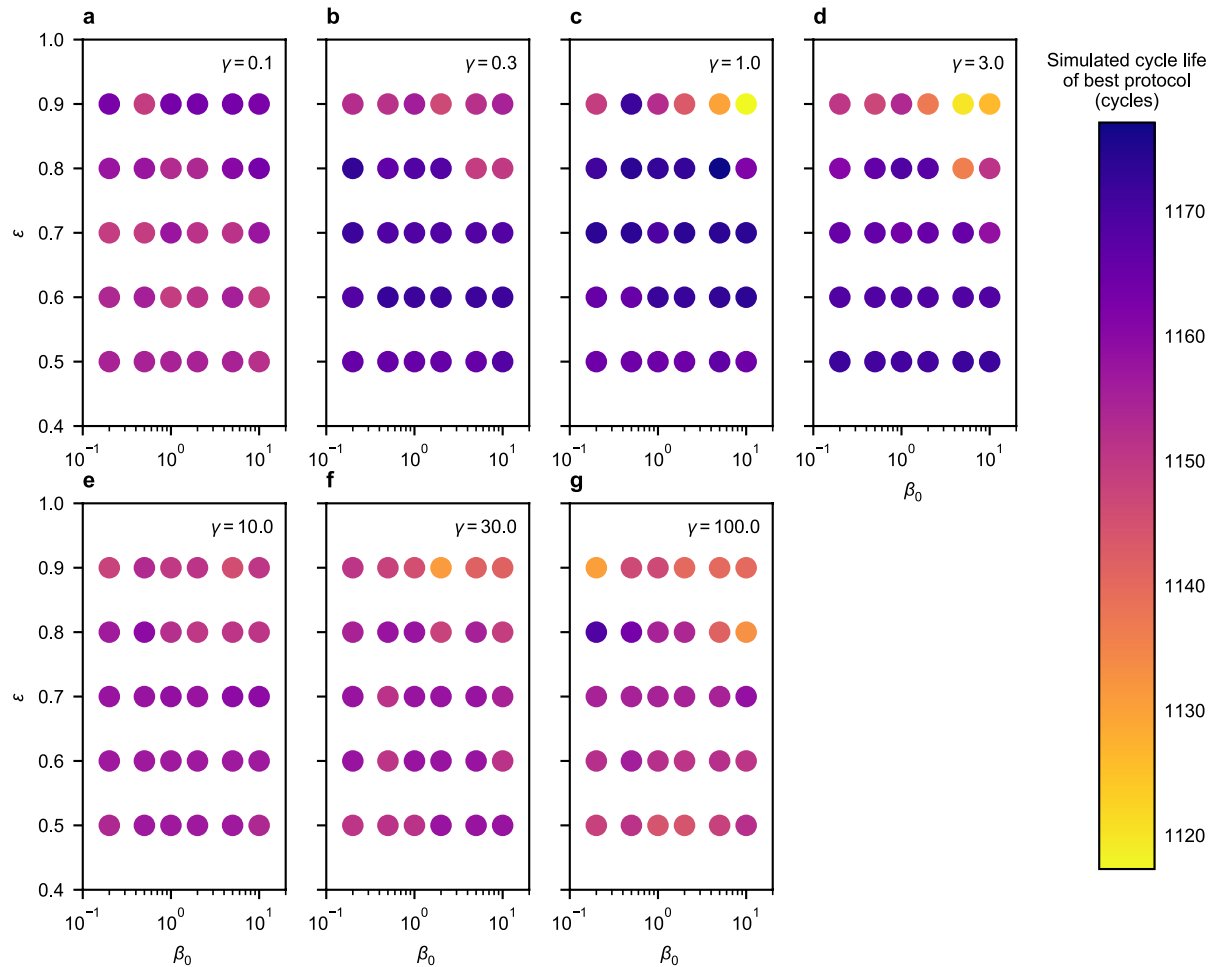
compares cycle lives on a battery (cell) basis; and the third column (**c**, **f** and **i**) compares the predicted ranking by cycle life via each method. Orange points represent the top three CLO-estimated protocols, blue points represent protocols inspired by the battery literature (Methods), and green points represent protocols selected to sample the distribution of estimated cycle lives. The error bars represent the 95% confidence intervals ($n = 5$; $n = 4$ for the early predictions of 3.6C-6.0C-5.6C-4.755C). The Pearson correlation coefficient and Kendall rank correlation coefficients are displayed for all relevant cycle life and ranking plots, respectively.



Extended Data Fig. 8 | Closed-loop performance under resource constraints.

Comparison of the closed loop with and without the Bayesian optimization algorithm (that is, with and without the explore/exploit component) as a function of number of channels and number of rounds in the 224-protocol space, using the first-principles simulator as the ground-truth source for cycle lives. Early prediction is not included. Each point represents the mean of 100

simulations; error bars represent the 95% confidence intervals ($n=100$). Early prediction is not incorporated into these simulations. The complete closed loop (that is, with Bayesian optimization) consistently outperforms the closed loop without Bayesian optimization. Bayesian optimization offers the largest advantage when the number of channels is low relative to the number of protocols.



Extended Data Fig. 9 | Hyperparameter sensitivity analysis on a cycle life simulator. The true cycle life of the best charging protocol as estimated by CLO, averaged over ten random seeds, is plotted as a function of the initial exploration constant (β_0), the exploration decay factor (ϵ) and the kernel bandwidth (γ). The values of all other hyperparameters are consistent with the values indicated in the ‘BO hyperparameter optimization’ Methods section and

in Supplementary Table 5. Overall, CLO achieves acceptable performance over a range of hyperparameter combinations; the highest-cycle-life protocols as estimated by the best and worst hyperparameter combinations differ by only 60 cycles. In our real-world CLO experiment, the selected hyperparameters are $\beta_0 = 5.0$, $\epsilon = 0.5$ and $\gamma = 1$; this combination performed well on a variety of simulated parameter spaces and budgets.

Extended Data Table 1 | Selected charging protocols for validation

Charging protocol	CLO-estimated cycle life	Early-predicted cycle life (from validation)	Final cycle life (from validation)	Source
3.6C-6.0C-5.6C-4.755C	1103 ± 131	1013 ± 115	755 ± 81	Zhang ³⁹
4.4C-5.6C-5.2C-4.252C	1174 ± 76	1056 ± 127	884 ± 132	Protocol with third-highest CLO-estimated mean cycle life
4.8C-5.2C-5.2C-4.160C	1185 ± 78	1047 ± 49	890 ± 90	Protocol with highest CLO-estimated mean cycle life
5.2C-5.2C-4.8C-4.160C	1183 ± 86	1098 ± 134	912 ± 118	Protocol with second-highest CLO-estimated mean cycle life
6.0C-5.6C-4.4C-3.834C	954 ± 164	963 ± 26	880 ± 85	
7.0C-4.8C-4.8C-3.652C	876 ± 183	964 ± 43	870 ± 70	Samsung patents ^{40,41}
8.0C-4.4C-4.4C-3.940C	818 ± 212	854 ± 44	702 ± 51	Notten <i>et al.</i> ⁴²
8.0C-6.0C-4.8C-3.000C	775 ± 273	698 ± 40	584 ± 60	Tesla patents ^{43,44}
8.0C-7.0C-5.2C-2.680C	648 ± 174	580 ± 68	496 ± 49	

The columns represent the CLO-estimated mean cycle lives of each protocol, early predictions in the validation experiment and the final tested cycle lives. For the CLO-estimated cycle lives, the errors represent the CLO-estimated standard deviation after round 4 (σ_{rd4}). For the early-predicted and final cycle lives from validation, the errors represent 95% confidence intervals ($n = 5$; but $n = 4$ for the early predictions of 3.6C-6.0C-5.6C-4.755C). The two protocols without a source were selected to obtain a representative sampling from the distribution of CLO-estimated cycle lives. Literature fast-charging protocols are from refs. ^{39–44}.

Synthesis of rare sugar isomers through site-selective epimerization

<https://doi.org/10.1038/s41586-020-1937-1>

Yong Wang¹, Hayden M. Carder¹ & Alison E. Wendlandt^{1✉}

Received: 11 November 2019

Accepted: 3 January 2020

Published online: 15 January 2020

Glycans have diverse physiological functions, ranging from energy storage and structural integrity to cell signalling and the regulation of intracellular processes¹. Although biomass-derived carbohydrates (such as D-glucose, D-xylose and D-galactose) are extracted on commercial scales, and serve as renewable chemical feedstocks and building blocks^{2,3}, there are hundreds of distinct monosaccharides that typically cannot be isolated from their natural sources and must instead be prepared through multistep chemical or enzymatic syntheses^{4,5}. These ‘rare’ sugars feature prominently in bioactive natural products and pharmaceuticals, including antiviral, antibacterial, anticancer and cardiac drugs^{6,7}. Here we report the preparation of rare sugar isomers directly from biomass carbohydrates through site-selective epimerization reactions. Mechanistic studies establish that these reactions proceed under kinetic control, through sequential steps of hydrogen-atom abstraction and hydrogen-atom donation mediated by two distinct catalysts. This synthetic strategy provides concise and potentially extensive access to this valuable class of natural compounds.

Simple structural and storage polymers including starch, cellulose, and hemicellulose are important sources of the monosaccharide feedstocks D-glucose, D-xylose, D-galactose, D-mannose and L-arabinose (Fig. 1a). Isomerization is a useful strategy for the synthesis of so-called ‘rare’ sugars (Fig. 1b) from biomass precursors; however, these processes remain challenging owing to the structural and stereochemical complexity of sugars. Chemical isomerization reactions (for example, the Lobry de Bruyn–Alberda van Ekenstein and Bilik reactions) are typically unselective, leading to complex thermodynamic distributions of products and often intractable separations (Extended Data Fig. 1a)⁸. In contrast, enzymatic methods offer an added level of precision and have emerged as a powerful synthetic alternative to chemical strategies⁹. Enzymatic isomerizations feature prominently in industrial sugar processing, including in the syntheses of D-fructose and D-ribose from D-glucose^{10,11}. In principle, multistep enzymatic synthesis also provides synthetic access to rare hexose, pentose and tetrose isomers but low yields and prohibitive production costs nonetheless limit implementation of these strategies^{12,13}. For example, D-allose has potential value as a low-glycemic sweetener and shows promising anti-inflammatory and immunosuppressive activity¹⁴. The enzymatic synthesis of D-allose can be achieved in overall 2.5% yield from D-glucose through sequential treatment with D-xylose isomerase (50% yield), D-tagatose 3-epimerase (20% yield), and L-rhamnose isomerase (25% yield allose + 8% yield altrose) (Extended Data Fig. 1b)¹⁵. Like chemical isomerizations, nearly all enzymatic isomerizations proceed through reversible polar enolization mechanisms under equilibrium control. Maximum product yields are therefore dictated by thermodynamic considerations under reaction conditions constrained by temperature-dependent enzyme activity. Reaction scope is also mechanistically restricted: for example, 2-deoxygenated sugars cannot be substrates under enolization-based isomerization conditions.

We envisioned that rare sugar isomers could be obtained directly from biomass-derived carbohydrates through site-selective radical epimerization reactions. A kinetically controlled epimerization process would require C–H bond cleavage and C–H bond formation steps to proceed through distinct mechanisms but could in principle afford product yields and selectivities exceeding those observed under canonical equilibrium-controlled sugar isomerization conditions. Our approach draws inspiration from recent reports of enzymatic radical epimerizations mediated by the cofactor S-adenosyl methionine (SAM)^{16,17}. For example, in the biosynthesis of neomycin B from neomycin C, a high-energy 5′-deoxyadenosyl radical abstracts the C5 hydrogen atom from the terminal saccharide of neomycin C (Fig. 1c). Subsequent re-delivery of a hydrogen atom to the opposite face is achieved via a pendant cysteine thiol. Within the enzyme active site, diastereoselective hydrogen-atom transfer (HAT) steps are thus promoted through exquisite spatial control over the hydrogen-atom donor/acceptor pairs. The architecture of the enzyme further mitigates any chemical incompatibility of the 5′-deoxyadenosyl radical and cysteine thiol. Outside the physical context of an enzyme, however, other interactions would be necessary to achieve kinetic control and prevent reagent donor/acceptor quenching. Although several notable synthetic HAT-mediated epimerization reactions have been developed, these methods almost universally proceed through reversible HAT steps, affording equilibrium product distributions¹⁸. However, a deracemization reaction recently reported by Knowles has established the conceptual viability of contra-thermodynamic radical isomerization through sequential proton-coupled electron transfer and HAT steps¹⁹.

In addition to kinetic challenges posed by employing chemically incompatible reagents, the success of this transformation further requires numerous nearly identical C–H bonds to be distinguished within the context of a highly polar, densely functionalized,

¹Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, USA. ✉e-mail: awendlan@mit.edu

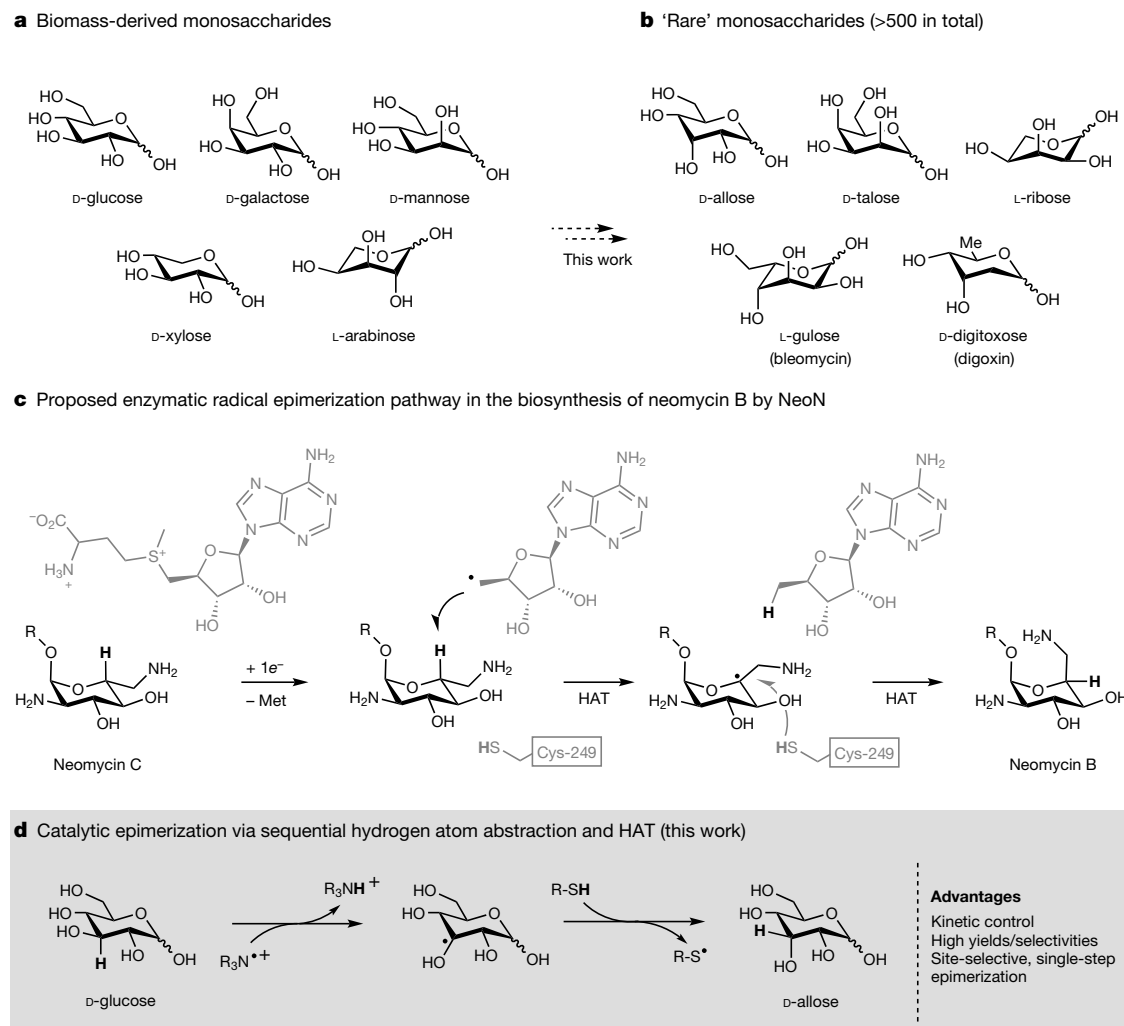


Fig. 1 | Approaches to the epimerization of sugars. **a**, Only a handful of monosaccharides can be obtained from natural sources. **b**, Hundreds of rare sugars exist and feature prominently in glycosylated natural products. **c**, Radical SAM-dependent epimerization by NeoN of neomycin C involves

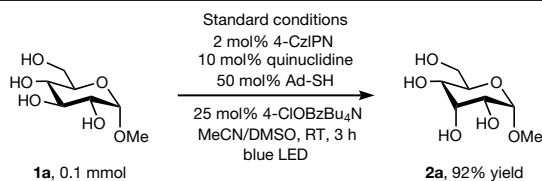
sequential HAT from substrate to the 5'-deoxyadenoxyl radical, and from cysteine thiol (Cys-249-SH) to the substrate. (Met, methionine.) **d**, Proposed direct chemical epimerization of biomass sugars to the rare sugars described here.

stereochemically complex unprotected carbohydrate substrate. Indeed, while site-selective functionalization of carbohydrate O-H bonds has been the subject of considerable recent attention²⁰, examples of selective carbohydrate C-H oxidation and functionalization are extremely limited²¹. Precedents set by the Minnaard, Waymouth and Muramatsu laboratories have established the feasibility of site-selective oxidation of minimally-protected monosaccharides to keto-sugars using Pd(II)/benzoquinone, Pd(II)/O₂, and Sn(IV)/Br₂ catalyst systems, respectively^{22–24}. Recent work by Minnaard has demonstrated that diastereoselective C-H alkylation of glucose derivatives can be achieved through site-selective hydrogen-atom abstraction by a quinuclidinium radical cation generated under photoredox conditions²⁵. Taylor has further expanded the scope of this transformation by employing a borinic acid co-catalyst to promote stereoretentive C-H alkylation of *cis*-diol-containing monosaccharides²⁶. Building on these findings, here we report a strategy for the synthesis of rare sugar isomers from biomass-derived precursors through the site-selective epimerization of unprotected sugars and glycans (Fig. 1d).

After extensive exploration of reaction conditions, the minimally protected model substrate, α -methylglucose, was found to react under photochemical conditions to afford α -methylallose as the sole reaction product in 92% yield within 3 h (Fig. 2). Optimal reaction conditions employ catalytic quantities of

1,2,3,5-tetrakis(carbazol-9-yl)-4,6-dicyanobenzene (4-CzIPN), quinuclidine, adamantane thiol and tetrabutylammonium *p*-chlorobenzoate in MeCN/DMSO at room temperature under blue light irradiation (see Supplementary Fig. 1). No epimerization was observed to occur in the absence of photocatalyst, thiol or light, and only trace product (<1%) was observed in the absence of quinuclidine. The reaction yield was considerably diminished in the absence of benzoate base (29% yield) or by employing tetrabutylammonium benzoate as the base (63% yield). Ir[(dF(CF₃))ppy]₂(dtbpy)]PF₆ (IrF) is also an effective photoredox catalyst for this transformation, promoting the desired reaction at only 1 mol% loading. However, the observation of oxidation products, as well as considerable cost differences (\$5 per mmol for 4-CzIPN versus \$935 per mmol for IrF; dollar prices are in US\$ in 2020) led us to select 4-CzIPN as the preferred reagent. Alkyl thiols were uniquely effective at promoting epimerization: no reaction was observed using thiophenols or thiobenzoic/thioacetic acid derivatives, nor when using any other hydrogen-atom donor surveyed (see Supplementary Information section 8).

Although the basis for site-selectivity is not yet fully understood, the C3 selectivity observed here is congruent with the substrate-controlled selectivity previously noted by both Waymouth and Minnaard in oxidation reactions^{22,23,25}. Nuclear magnetic resonance (NMR) titration experiments reveal an equilibrium interaction between **1a** and



Entry	Variation from standard conditions	Percentage yield (% RSM)
1	No photocatalyst	0% (99%)
2	No quinuclidine	< 1% (99%)
3	No thiol	0% (98%)
4	No blue LED	0% (99%)
5	No 4-ClOBzBu ₄ N	29% (71%)
6	Bu ₄ NOBz instead of 4-ClOBzBu ₄ N	63% (35%)
7	1% Ir[(dF(CF ₃)ppy) ₂ (dtbpy)]PF ₆ instead of 4-CzIPN	88% (2%)
8	DABCO instead of quinuclidine	0% (99%)
9	Hydrogen atom source ^a	0% (99%)

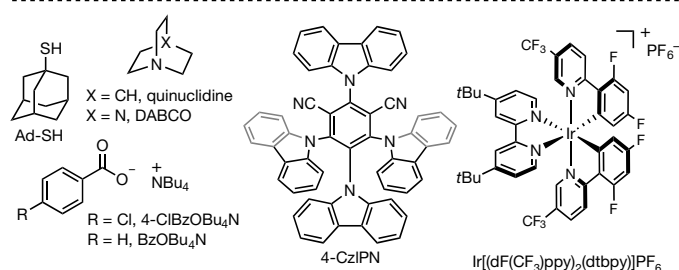


Fig. 2 | Epimerization of α -methylglucose to α -methylallose. Effect of changes to optimized reaction conditions. Yields determined by proton nuclear magnetic resonance (¹H NMR) analysis using 4-fluoroanisole as internal standard. RSM, recovered starting material; MeCN, acetonitrile; DMSO, dimethylsulfoxide; RT, room temperature; LED, light-emitting diode; Me, methyl; DABCO, 1,4-diazabicyclo[2.2.2]octane; Ad-SH, adamantane thiol; Bz, benzoyl. ^aSee Supplementary Information sections 5 and 8 for full experimental details.

tetrabutylammonium p-chlorobenzoate with alteration of substrate ¹J_(C-H) coupling constants, implicating the presence of hydrogen-bonding interactions between **1a** and the base with concomitant weakening of α C-H bonds²⁷. In contrast, no such interaction is observed between pentamethylated glucose and base, and importantly, no epimerization products were detected in reactions employing permethylated or peracetylated substrates.

A range of biomass-derived and abundant monosaccharides were evaluated as substrates under the optimized reaction conditions (Fig. 3). This strategy provides synthetic access to 4 of the 5 rare hexose isomers. In addition to D-allose products obtained from glucose-configured substrates, D-talose (**2b**), D-gulose (**2c**) and D-altrose (**2d**) products are obtained selectively from the reaction of β -methylgalactose, anhydrogalactose and anhydromannose, respectively. The biomass-derived pentose sugars D- α -methylxylose (**1f**) and L- β -methylarabinose (**1e**) afforded D- and L-ribose derivatives through C3 and C2 epimerization, respectively. Although D-ribose is accessible on a large scale through glucose fermentation, synthetic access to L-ribose remains extremely limited^{2,11}. Despite the presence of an electron-rich acetamide substituent, the N-acetylglucosamine derivative **1g** also undergoes productive epimerization, reacting to afford a 1.5:1 mixture of the C3- and C4-epimeric products, N-acetylallosamine, **2g**, and N-acetylgalactosamine, **3g**, respectively.

Completely unprotected monosaccharides also undergo selective epimerization under these conditions: for example, 42% yield D-allose **2k** is obtained from D-glucose (compared with a 2.5% total yield over

4 steps, using enzymatic methods), and 55% yield L-6-deoxytalose **2m** is obtained from the reaction of L-fucose. The reaction of D-2-deoxyglucose affords D-2-deoxyallose in 61% yield: importantly, epimerization of 2-deoxygenated sugars cannot be realized using alternative, enolization-based isomerization methods.

More complex glycans were subsequently evaluated to assess further the selectivity and functional group compatibility of the reaction conditions. Allosucrose, **2o**, can be obtained selectively from sucrose in 68% yield, and despite the presence of 14 stereogenic centres, the reaction of raffinose, **1p**, affords a singly epimerized product, **2p**, in 44% yield. Pyrimidine-containing pyranonucleoside, **1q**, reacted to afford the C3-epimeric product, **2q**. Finally, the C-glycoside SGLT2 inhibitor empagliflozin was also examined as a substrate, and was found to afford the C3-epimeric product, **2r**, in 42% yield with no observation of epimerization at any other position in the molecule.

We performed a series of experiments to gather insight into the underlying mechanism of this transformation. Re-subjecting 6-deoxy- β -methyltalose, **2h**, (formed in 70% yield from β -methylfucose, **1h**) to standard reaction conditions did not result in the formation of any β -methylfucose starting material (Fig. 4a). A similar experiment was conducted using α -methylallose and adamantane thiol-*d*1 (where '*d*1' indicates deuterium isotopic substitution). After 16 hours, 95% α -methylallose was recovered with 39% deuterium incorporation at the C3 position; no glucose products were detected. This experiment demonstrates that both α -methylglucose and α -methylallose can undergo hydrogen-atom abstraction but that both converge to the α -methylallose product (Fig. 4b). Taken in conjunction with established thermochemical data, these experiments provide preliminary evidence that these transformations do not proceed under simple equilibrium control²⁸.

To explore further the individual elementary steps of this reaction, Stern–Volmer fluorescence quenching was examined under two different sets of conditions (see Supplementary Information section 9). Preliminary experiments reveal that quinuclidine efficiently quenches the photocatalyst excited state, whereas adamantane thiol does not quench under the conditions examined. In the presence of both quinuclidine and adamantane thiol, quenching kinetics are identical to the 'quinuclidine only' conditions. These findings support a mechanism in which the excited photocatalyst is quenched by quinuclidine to generate quinuclidinium radical cation²⁹. When catalyst loading is increased to 3 mol% and the reaction is run in the absence of thiol, small quantities (3% yield) of 3-keto sugar are obtained (Fig. 4c). Importantly, no epimerization is observed under these conditions, nor under any conditions tested where alkyl thiol is not present in the reaction mixture (see Supplementary Information section 9). These experiments establish that photocatalyst and quinuclidine are sufficient for C-H cleavage to occur but insufficient for epimerization. Together, they support a mechanism in which the quinuclidinium radical cation mediates an irreversible hydrogen-atom abstraction step.

To probe the role of the thiol co-catalyst, an analogous set of fluorescence quenching experiments was carried out using 4-bromothiophenol in the place of adamantane thiol. Under these reaction conditions no epimerization or consumption of α -methylglucose is observed. As with adamantane thiol, minimal fluorescence quenching was observed using 4-bromothiophenol alone. However, in the presence of both 4-bromothiophenol and quinuclidine, a substantial increase in fluorescence quenching was observed relative to the case with quinuclidine alone (see Supplementary Information section 9). We postulated that this enhanced fluorescence quenching might be due to the oxidation of thiolate—generated in situ by deprotonation of thiol by quinuclidine, or through proton-coupled electron transfer from a quinuclidine/thiol hydrogen-bonded complex—to the corresponding thiyl radical. Indeed, sodium thiophenolate was also found to quench the photocatalyst at a tenfold-higher rate than quinuclidine alone, and NMR titration studies performed under comparable conditions identified a small equilibrium

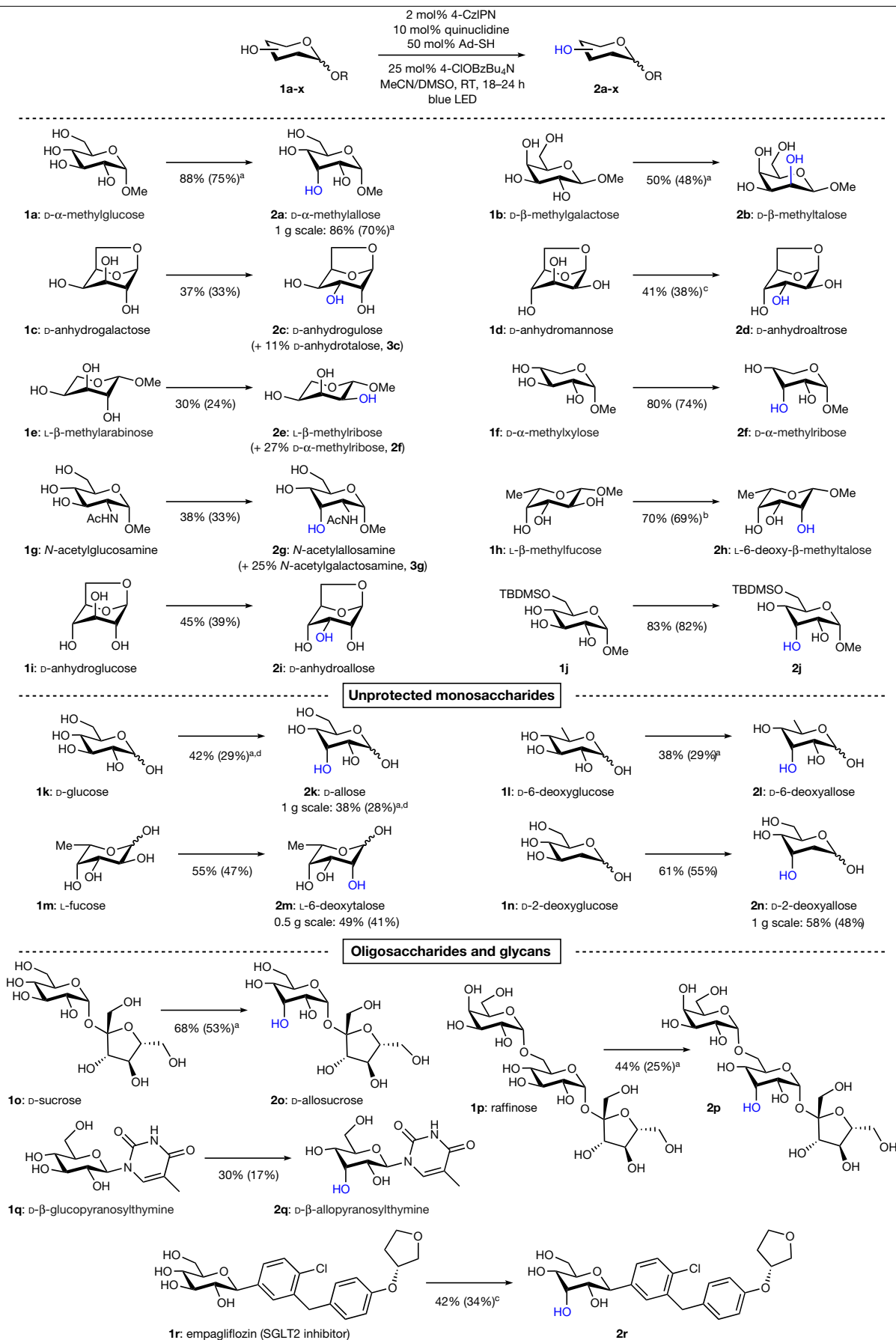


Fig. 3 | Site-selective epimerization of saccharides and glycans. Scope of the reaction with respect to substrate. Reactions were conducted at 0.3-mmol scale in duplicate; the average ^1H NMR yield is reported; duplicate reactions were combined, then isolated; isolated yields are in parentheses. ^a, Isolated in **406** | Nature | Vol 578 | 20 February 2020

peracetylated form. ^b, 1 mol% Ir[(dF(CF₃)ppy)₂(dtbpy)]PF₆ used instead of 4-CzIPN. ^c, Reaction was conducted using 5 mol% quinuclidine. ^d, Reaction was conducted using 2.5 mol% quinuclidine. See Supplementary Information Sections 5–7 for full experimental details and conditions.

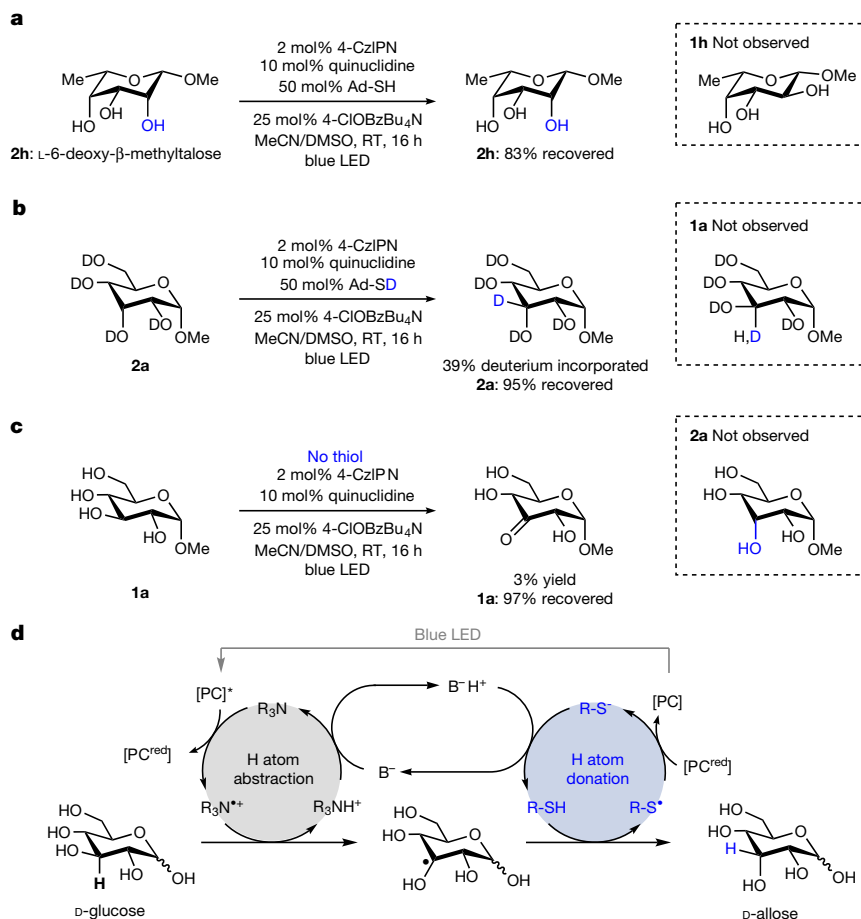


Fig. 4 | Mechanistic studies and proposed mechanism. a, No epimerization is observed when reaction products are re-subjected to the standard reaction conditions. **b**, Deuterium labelling studies indicate that the reaction product reacts under standard reaction conditions, but both epimers converge to a

common product. **c**, Reaction in the absence of thiol donor affords no isomerization product, implicating irreversible hydrogen-atom abstraction by the quinuclidinium radical cation. **d**, Plausible mechanistic pathway for selective isomerization reactions.

interaction ($K_{eq} = 42 \text{ M}^{-1}$) between 4-bromothiophenol ($pK_a = 9.0$ in DMSO) and quinuclidine (quinuclidinium conjugate acid, $pK_a = 9.8$ in DMSO), supporting the formation of thiolate in solution.

These experiments indicate that thiol acidity is an important parameter distinguishing productive versus unproductive reaction conditions. In the presence of quinuclidine, acidic thiols can be deprotonated to form thiolate salts. Preferential thiolate quenching of the photocatalyst results in the formation of thiyl radicals. No epimerization was observed under these, or any other photo-oxidative, photo-reductive and thermal conditions that we explored for the in situ generation of thiyl radical (see Supplementary Information section 9). These findings suggest that thiyl radical is not competent for hydrogen-atom abstraction. Accordingly, the thiol co-catalyst is instead implicated in a subsequent, irreversible HAT to the incipient sugar radical.

Collectively, the mechanistic studies presented here support a non-equilibrium epimerization mechanism proceeding through two sequential and distinct HAT steps: hydrogen-atom abstraction by quinuclidinium radical cation (homolytic bond dissociation enthalpy = $100 \text{ kcal mol}^{-1}$) from substrate, followed by HAT from thiol (87 kcal mol^{-1}) to the incipient sugar radical (Fig. 4d). Although both substrate and product can undergo hydrogen-atom abstraction by quinuclidinium radical cation, mechanistic data are consistent with irreversible hydrogen-atom abstraction followed by diastereoselective HAT from thiol. Attendant to a kinetically controlled epimerization mechanism, the reaction yields and product selectivities presented here exceed nearly all other direct isomerization yields reported so far, which have reflected thermodynamic product distributions.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-1937-1>.

- Varki, A. et al. (eds) *Essentials of Glycobiology* 3rd edn (Cold Spring Harbor Laboratory Press, 2017).
- de Leder Kremer, R. M. & Gallo-Rodriguez, C. Naturally occurring monosaccharides: properties and synthesis. *Adv. Carbohydr. Chem. Biochem.* **59**, 9–67 (2004).
- Chheda, J. N., Huber, G. W. & Dumesic, J. A. Liquid-phase catalytic processing of biomass-derived oxygenated hydrocarbons to fuels and chemicals. *Angew. Chem. Int. Ed.* **46**, 7164–7183 (2007).
- Imperiali, B. Bacterial carbohydrate diversity—a brave new world. *Curr. Opin. Chem. Biol.* **53**, 1–8 (2019).
- Herget, S. et al. Statistical analysis of the Bacterial Carbohydrate Structure Data Base (BCSDB): characteristics and diversity of bacterial carbohydrates in comparison with mammalian glycans. *BMC Struct. Biol.* **8**, 35 (2008).
- Frihed, T. G., Bols, M. & Pedersen, C. M. Synthesis of L-hexoses. *Chem. Rev.* **115**, 3615–3676 (2015).
- Elshahawi, S. I., Shaaban, K. A., Kharel, M. K. & Thorson, J. S. A comprehensive review of glycosylated bacterial natural products. *Chem. Soc. Rev.* **44**, 7591–7697 (2015).
- Angyal, S. The Lobry de Bruyn–Alberda van Ekenstein transformation and related reactions. In *Glycoscience: Epimerisation, Isomerisation and Rearrangement Reactions of Carbohydrates* (ed. Stütz, A. E.) 1–14 (Springer, 2001).
- Buchholz, K. & Seibel, J. Industrial carbohydrate biotransformations. *Carb. Res.* **343**, 1966–1979 (2008).
- Bhosale, S. H., Rao, M. B. & Deshpande, V. V. Molecular and industrial aspects of glucose isomerase. *Microbiol. Rev.* **60**, 280–300 (1996).
- Wulf, P. & Vandamme, E. Production of D-ribose by fermentation. *Appl. Microbiol. Biotechnol.* **48**, 141–148 (1997).

12. Granström, T. B., Takata, G., Tokuda, M. & Izumori, K. Izumoring: a novel and complete strategy for the bioproduction of rare sugars. *J. Biosci. Bioeng.* **97**, 89–94 (2004).
13. Zhang, W., Zhang, T., Jiang, B. & Mu, W. Enzymatic approaches to rare sugar production. *Biotechnol. Adv.* **35**, 267–274 (2017).
14. Hossain, M. A. et al. Effect of the immunosuppressants FK506 and D-allose on allogenic orthotopic liver transplantation in rats. *Transplant. Proc.* **32**, 2021–2023 (2000).
15. Menavvuri, B. T. et al. Efficient biosynthesis of D-allose from D-psicose by cross-linked recombinant L-rhamnose isomerase: separation of product by ethanol crystallization. *J. Biosci. Bioeng.* **101**, 340–345 (2006).
16. Kudo, F., Hoshi, S., Kawashima, T., Kamachi, T. & Eguchi, T. Characterization of a radical S-adenosyl-L-methionine epimerase, NeoN, in the last step of neomycin B biosynthesis. *J. Am. Chem. Soc.* **136**, 13909–13915 (2014).
17. Benjdia, A., Guillot, A., Ruffie, P., Leprince, J. & Berteau, O. Post-translational modification of ribosomally synthesized peptides by a radical SAM epimerase in *Bacillus subtilis*. *Nat. Chem.* **9**, 698–707 (2017).
18. Wang, Y. et al. Epimerization of tertiary carbon centers via reversible radical cleavage of unactivated C(sp³)–H bonds. *J. Am. Chem. Soc.* **140**, 9678–9684 (2018).
19. Shin, N. Y., Ryss, J. M., Zhang, X., Miller, S. J. & Knowles, R. R. Light-driven deracemization enabled by excited-state electron transfer. *Science* **366**, 364–369 (2019).
20. Dimakos, V. & Taylor, M. S. Site-selective functionalization of hydroxyl groups in carbohydrate derivatives. *Chem. Rev.* **118**, 11457–11517 (2018).
21. Frihed, T. G., Bols, M. & Pedersen, C. M. C–H functionalization on carbohydrates. *Eur. J. Org. Chem.* 2740–2756 (2016).
22. Jäger, M., Hartmann, M., de Vries, J. G. & Minnaard, A. J. Catalytic regioselective oxidation of glycosides. *Angew. Chem. Int. Ed.* **52**, 7809–7812 (2013).
23. Chung, K. & Waymouth, R. M. Selective catalytic oxidation of unprotected carbohydrates. *ACS Catal.* **6**, 4653–4659 (2016).
24. Muramatsu, W. Catalytic and regioselective oxidation of carbohydrates to synthesize keto-sugars under mild conditions. *Org. Lett.* **16**, 4846–4849 (2014).
25. Wan, I. C., Witte, M. D. & Minnaard, A. J. Site-selective carbon–carbon bond formation in unprotected monosaccharides using photoredox catalysis. *Chem. Commun.* **53**, 4926–4929 (2017).
26. Dimakos, V., Su, H. Y., Garrett, G. E. & Taylor, M. S. Site-selective and stereoselective C–H alkylations of carbohydrates via combined diarylborinic acid and photoredox catalysis. *J. Am. Chem. Soc.* **141**, 5149–5153 (2019).
27. Maiti, N. C., Zhu, Y., Carmichael, I., Serianni, A. S. & Anderson, A. E. ¹J_{CH} correlates with alcohol hydrogen bond strength. *J. Org. Chem.* **71**, 2878–2880 (2006).
28. Goldberg, R. N. & Tewari, Y. B. Thermodynamic and transport properties of carbohydrates and their monophosphates: the pentoses and hexoses. *J. Phys. Chem. Ref. Data* **18**, 809–880 (1989).
29. Jeffrey, J. L., Terrett, J. A. & Macmillan, D. W. C. O–H hydrogen bonding promotes H atom transfer from C–H bonds for C–alkylation of alcohols. *Science* **349**, 1532–1536 (2015).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Data availability

All data supporting the findings of this paper are available within the Article and its supplementary information files.

Acknowledgements We thank A. Seim for checking the reaction procedure and X. Gu for help with substrate synthesis. A.E.W. also thanks E. Kwan, A. Radosevich, D. Suess and Z. Wickens for discussions. Financial support for this work was provided by the Massachusetts Institute of Technology and the National Science Foundation (NSF) for funding through the National Science Foundation Graduate Research Fellowships Program (NSF-GRFP) to H.M.C.

Author contributions A.E.W. and Y.W. conceived the work. A.E.W., Y.W. and H.M.C. designed the experiments. Y.W. and H.M.C. conducted the experiments. A.E.W. directed the research and wrote the manuscript with input from all authors.

Competing interests The authors declare no competing interests.

Additional information

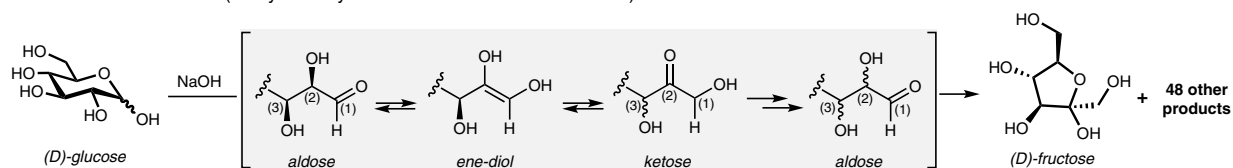
Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-1937-1>.

Correspondence and requests for materials should be addressed to A.E.W.

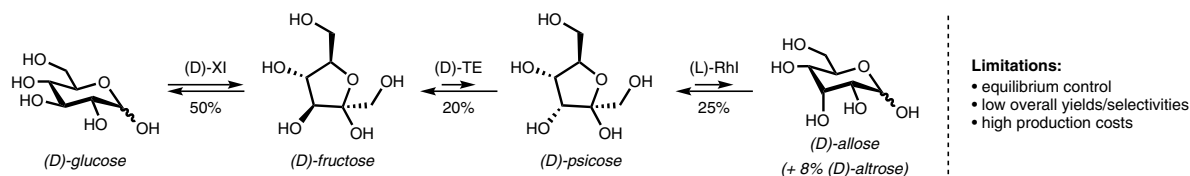
Peer review information *Nature* thanks Christian Marcus Pedersen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

a Chemical isomerization (Lobry de Bruyn-Alberda van Ekenstein Reaction)



b "Izumoring" enzymatic isomerization



Limitations:

- equilibrium control
- low overall yields/selectivities
- high production costs

Extended Data Fig. 1 | Chemical and enzymatic isomerizations proceed through polar aldose-ketose mechanisms. a, Chemical isomerization reactions of glucose lead to unselective and complex thermodynamic distribution of products. **b,** 'Izumoring' enzymatic synthesis of D-allose

proceeds through reversible polar enolization mechanisms under equilibrium control; see ref.¹². D-XI is D-xylose isomerase; D-TE is D-tagatose 3-epimerase; and L-Rhl is L-rhamnose isomerase.

Preindustrial $^{14}\text{CH}_4$ indicates greater anthropogenic fossil CH_4 emissions

<https://doi.org/10.1038/s41586-020-1991-8>

Received: 27 May 2019

Accepted: 27 November 2019

Published online: 19 February 2020

 Check for updates

Benjamin Hmiel^{1✉}, V. V. Petrenko¹, M. N. Dyonisius¹, C. Buizert², A. M. Smith³, P. F. Place¹, C. Harth⁴, R. Beaudette⁴, Q. Hua³, B. Yang³, I. Vimont⁵, S. E. Michel⁶, J. P. Severinghaus⁴, D. Etheridge⁷, T. Bromley⁸, J. Schmitt⁹, X. Fain¹⁰, R. F. Weiss⁴ & E. Dlugokencky¹¹

Atmospheric methane (CH_4) is a potent greenhouse gas, and its mole fraction has more than doubled since the preindustrial era¹. Fossil fuel extraction and use are among the largest anthropogenic sources of CH_4 emissions, but the precise magnitude of these contributions is a subject of debate^{2,3}. Carbon-14 in CH_4 ($^{14}\text{CH}_4$) can be used to distinguish between fossil (^{14}C -free) CH_4 emissions and contemporaneous biogenic sources; however, poorly constrained direct $^{14}\text{CH}_4$ emissions from nuclear reactors have complicated this approach since the middle of the 20th century^{4,5}. Moreover, the partitioning of total fossil CH_4 emissions (presently 172 to 195 teragrams CH_4 per year)^{2,3} between anthropogenic and natural geological sources (such as seeps and mud volcanoes) is under debate; emission inventories suggest that the latter account for about 40 to 60 teragrams CH_4 per year^{6,7}. Geological emissions were less than 15.4 teragrams CH_4 per year at the end of the Pleistocene, about 11,600 years ago⁸, but that period is an imperfect analogue for present-day emissions owing to the large terrestrial ice sheet cover, lower sea level and extensive permafrost. Here we use preindustrial-era ice core $^{14}\text{CH}_4$ measurements to show that natural geological CH_4 emissions to the atmosphere were about 1.6 teragrams CH_4 per year, with a maximum of 5.4 teragrams CH_4 per year (95 per cent confidence limit)—an order of magnitude lower than the currently used estimates. This result indicates that anthropogenic fossil CH_4 emissions are underestimated by about 38 to 58 teragrams CH_4 per year, or about 25 to 40 per cent of recent estimates. Our record highlights the human impact on the atmosphere and climate, provides a firm target for inventories of the global CH_4 budget, and will help to inform strategies for targeted emission reductions^{9,10}.

Atmospheric measurements of carbon-13 in methane ($\delta^{13}\text{CH}_4$) have been used to estimate the fossil fraction of the contemporaneous CH_4 budget³. This approach relies on having accurate estimates of the $\delta^{13}\text{C}$ signatures of the major CH_4 source categories (fossil, microbial and biomass burning) and the strength of the biomass burning source. Large uncertainties in these parameters in the past preclude accurate $\delta^{13}\text{CH}_4$ -based estimates of preindustrial-era fossil CH_4 emissions^{8,11,12}. Radiocarbon (^{14}C) is an ideal tracer for quantifying the fossil component of the atmospheric CH_4 budget because all ^{14}C in fossil CH_4 has decayed. By contrast, biogenic CH_4 sources (wetlands, biomass burning) have a ^{14}C activity similar to that of contemporaneous atmospheric CO_2 (ref. 4,8). Interpretation of atmospheric $^{14}\text{CH}_4$ measurements from 1987–2000 suggests that the fossil fraction of the contemporary CH_4 budget is

$30 \pm 2.3\%$ (ref. 13; 1σ). However, the interpretation of atmospheric $^{14}\text{CH}_4$ in recent decades has been complicated by (1) rapidly changing atmospheric $^{14}\text{CO}_2$ (from above-ground nuclear testing and fossil fuel emissions) that propagates into biospheric CH_4 emissions¹³, and (2) direct $^{14}\text{CH}_4$ emissions from nuclear power plants^{4,5}. By contrast, palaeoatmospheric $^{14}\text{CH}_4$ measurements from ice cores offer a direct constraint on natural geological CH_4 emissions without these complications. Whereas geological CH_4 emissions have the potential to change on tectonic- and glacial-cycle timescales¹⁴, they have very probably been constant over the past few centuries. The preindustrial-era emission estimates can therefore be applied to the modern CH_4 budget with confidence.

Ice core $^{14}\text{CH}_4$ analysis is challenging owing to both the very large sample requirement (~1,000 kg of ice) and interference from in situ

¹Department of Earth and Environmental Sciences, University of Rochester (UR), Rochester, NY, USA. ²College of Earth, Ocean and Atmospheric Sciences, Oregon State University (OSU), Corvallis, OR, USA. ³Australian Nuclear Science and Technology Organisation (ANSTO), Lucas Heights, New South Wales, Australia. ⁴Scripps Institution of Oceanography (SIO), University of California San Diego, La Jolla, CA, USA. ⁵Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado and National Oceanic and Atmospheric Administration (NOAA) Global Monitoring Division (GMD), Boulder, CO, USA. ⁶Institute of Arctic and Alpine Research (INSTAAR), University of Colorado, Boulder, CO, USA. ⁷Climate Science Centre, Commonwealth Scientific and Industrial Research Organisation (CSIRO) Oceans and Atmosphere, Aspendale, Victoria, Australia. ⁸National Institute of Water and Atmospheric Research (NIWA), Wellington, New Zealand. ⁹Climate and Environmental Physics, Physics Institute and Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland. ¹⁰University of Grenoble Alpes, CNRS, IRD, Grenoble INP, Institut des Géosciences de l'Environnement (IGE), Grenoble, France. ¹¹NOAA, Earth System Research Laboratory (ESRL), Boulder, CO, USA.

✉e-mail: bhmiel@ur.rochester.edu

cosmogenic ^{14}C production within the ice crystal lattice¹⁵. We address the former by using a large-diameter ice drill and a large-volume ice-melting apparatus (Supplementary Information section 1) to obtain sufficient CH_4 (~20 $\mu\text{g C}$) for ^{14}C analysis by accelerator mass spectrometry. To address the latter, we follow the established^{18,16} approach of analysing ^{14}C of carbon monoxide (CO) in parallel with $^{14}\text{CH}_4$. ^{14}CO is very sensitive to in situ cosmogenic ^{14}C production¹⁵ and can be used to precisely establish the effective cosmic ray exposure history of each sample. We then correct the $^{14}\text{CH}_4$ data using the known in situ cosmogenic $^{14}\text{CH}_4/^{14}\text{CO}$ production ratio in ice¹⁵ (Supplementary Information sections 5, 6). The in situ cosmogenic $^{14}\text{CH}_4$ component in the samples used in this study is much lower (<2% of total $^{14}\text{CH}_4$) than in ablation-zone ice used in previous palaeoatmospheric $^{14}\text{CH}_4$ studies (~30% of total $^{14}\text{CH}_4$)^{8,16}. We present new $^{14}\text{CH}_4$ data from large-volume ice core samples and firn air sampling from Summit, Greenland, which we combine with prior firn air $^{14}\text{CH}_4$ measurements from Law Dome DSSW20K⁴ and Megadunes¹⁷, Antarctica. Our combined record spans from about 1750 to 2013 and captures the evolution of atmospheric $^{14}\text{CH}_4$ since the preindustrial era (Fig. 1). The movement of gases within the firn and closure into bubbles is characterized using a firn air transport model¹⁸, and the time series of atmospheric $^{14}\text{CH}_4$ is reconstructed using a matrix inversion technique^{19,20} (Supplementary Information section 9).

Our atmospheric $^{14}\text{CH}_4$ reconstruction (Fig. 1) is indistinguishable from the $^{14}\text{CO}_2$ -derived contemporaneous biogenic $^{14}\text{CH}_4$ signature (blue curve, Supplementary Information section 10) before 1880, suggesting very low natural geological CH_4 emissions. Atmospheric $^{14}\text{CH}_4$ began to decrease around 1880, coincident with substantial increases in the use of coal, oil and natural gas (Fig. 2)²¹. The precise timing of the $^{14}\text{CH}_4$ minimum (in the 1940s in our reconstruction) is difficult to establish owing to the broad age distributions of individual firn air and ice core samples, as well as the smoothing applied by the matrix inversion technique to address the non-uniqueness of the solution¹⁹. Beyond this fossil ^{14}C minimum, our samples are affected by the propagation of ^{14}C from atmospheric nuclear testing into the carbon cycle²² and by emissions from nuclear power plants (starting in the 1970s), which drove a sustained $^{14}\text{CH}_4$ increase despite decreasing $^{14}\text{CO}_2$ ^{4,5}. We calculate the fossil CH_4 fraction and develop a time series of fossil CH_4 emissions (Fig. 2) using a one-box atmospheric model (Supplementary Information section 10). The broad age distributions of our air samples (Supplementary Fig. 3) result in a smoothed representation of the atmospheric $^{14}\text{CH}_4$ history that cannot capture the abrupt increase of bomb $^{14}\text{CO}_2$ (and subsequently $^{14}\text{CH}_4$) starting in 1955. Therefore, we interpret the fossil CH_4 fraction only before the 1940s. We find an increase in the total (geological plus anthropogenic) fossil emissions from negligible CH_4 emissions in the mid-19th century to 64.8 teragrams CH_4 per year ($\text{Tg CH}_4 \text{ yr}^{-1}$) in 1940.

Assuming that the oldest ice core $^{14}\text{CH}_4$ sample in our reconstruction (mean age 1756 AD; Fig. 1) is devoid of anthropogenic fossil CH_4 contributions, we use the contemporaneous biogenic $^{14}\text{CH}_4$ source signature to calculate the natural geological CH_4 emissions during the preindustrial era: 1.6 $\text{Tg CH}_4 \text{ yr}^{-1}$ with a 95% confidence interval (CI) maximum of 5.4 $\text{Tg CH}_4 \text{ yr}^{-1}$ (Supplementary Information section 10, Supplementary Fig. 5). Our 95% confidence limit of 5.4 $\text{Tg CH}_4 \text{ yr}^{-1}$ agrees well with, and provides a tighter constraint than, the only other published $^{14}\text{CH}_4$ -based estimate of natural geological CH_4 emissions from ice cores, which sampled air from the most recent deglaciation (0 to 15.4 $\text{Tg CH}_4 \text{ yr}^{-1}$, 95% CI range)⁸.

Our result is much lower than estimates from recent source inventory ('bottom-up') studies typically used in global CH_4 budgets², which suggest natural geological emissions of ~40–60 $\text{Tg CH}_4 \text{ yr}^{-1}$ (ref. 6). A recent study⁷ aimed at developing gridded maps of geological CH_4 emissions revised this estimate downwards to 37 $\text{Tg CH}_4 \text{ yr}^{-1}$ on the basis of data and modelling specifically targeted for gridding; however, the CH_4 emissions increased to 43–50 $\text{Tg CH}_4 \text{ yr}^{-1}$ when extrapolated to account for temporal variability in mud volcano eruptions and onshore

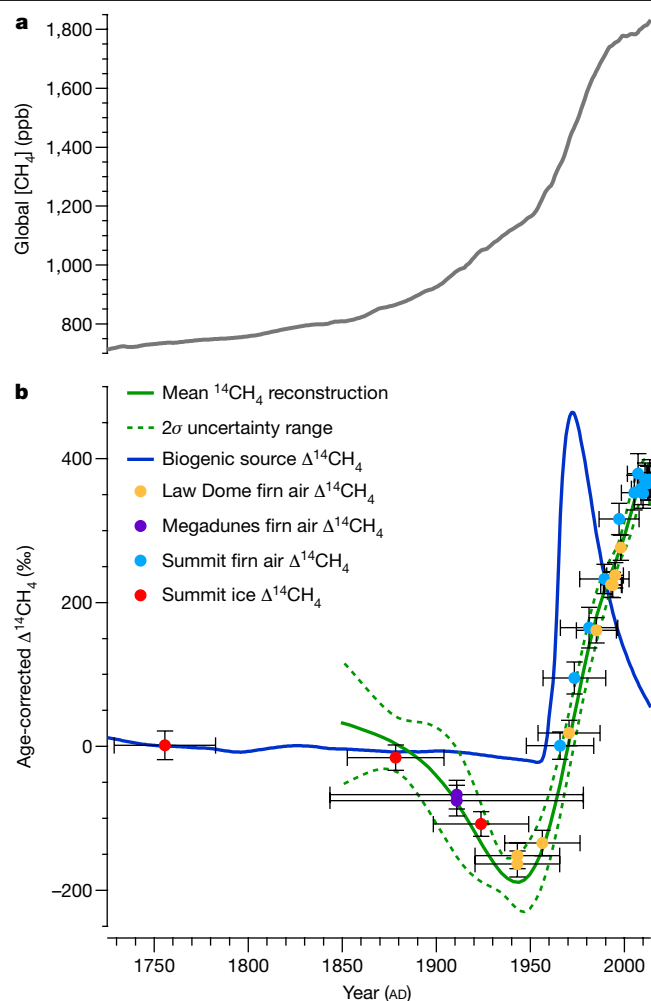


Fig. 1 | Reconstruction of atmospheric $^{14}\text{CH}_4$ from firn air and ice core data. **a**, Global CH_4 mole fraction, $[\text{CH}_4]$, reconstructed from ice core, firn air and atmospheric measurements¹, ppb, parts per billion. **b**, Reconstructed history of atmospheric $\Delta^{14}\text{CH}_4$ from firn air and ice core samples (this study). Dotted lines represent the 95% confidence range based on all calculated $^{14}\text{CH}_4$ histories using three different inversion methods (Supplementary Information section 9). Ice core and firn air $\Delta^{14}\text{CH}_4$ measurements are shown at the mean age of the modelled air age distribution. Vertical error bars on the $\Delta^{14}\text{CH}_4$ data from each site represent the 2σ uncertainty for each sample after corrections (Supplementary Information Tables 2, 6), and horizontal error bars represent $\pm 2\Delta$, where Δ is the spectral width of the sample-air age distribution²⁰. We also plot the $^{14}\text{CH}_4$ signature of the contemporaneous biogenic source (blue; Supplementary Information section 10). Our time series begins in 1850 because the age distributions of the collected ice core samples have poor coverage of air from ~1780 to 1850 (Supplementary Information section 10, Supplementary Fig. 3B).

or submarine geological seeps that lack location-specific measurements. Natural fossil CH_4 emissions of about 40 $\text{Tg CH}_4 \text{ yr}^{-1}$ (out of total preindustrial-era CH_4 emissions of 215 $\text{Tg CH}_4 \text{ yr}^{-1}$; Supplementary Fig. 5) would result in a preindustrial-era $\Delta^{14}\text{CH}_4$ of around ~185‰, which is in clear disagreement with our data ($1.5\% \pm 21.2\%$, 2σ ; Fig. 1). Bringing our ^{14}C results into agreement with the bottom-up estimates of natural fossil CH_4 emissions would require an order-of-magnitude larger correction for in situ cosmogenic $^{14}\text{CH}_4$. This would in turn require either an order-of-magnitude higher ^{14}CO content in the sampled ice or an order-of-magnitude higher in situ $^{14}\text{CH}_4/^{14}\text{CO}$ production ratio; both of these possibilities are well outside the respective uncertainties. The added uncertainties arising from the in situ and procedural corrections to the measured $^{14}\text{CH}_4$ are also too small to explain the disagreement

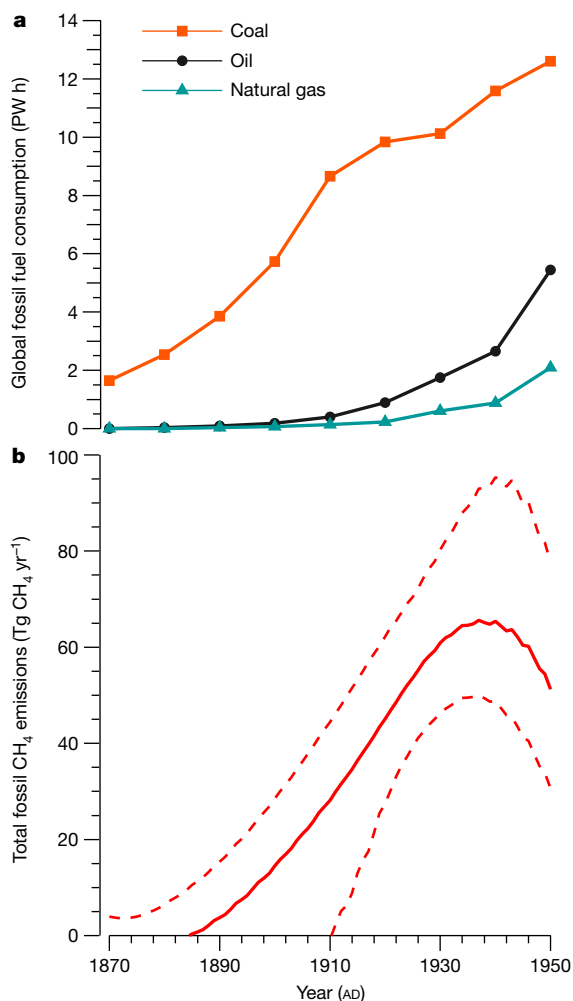


Fig. 2 | Growth in fossil CH₄ emissions and fossil fuel consumption. **a**, Historical fossil fuel energy consumption²¹. **b**, Calculated total fossil CH₄ emission history (solid line) from the one-box model (Supplementary Information section 10). The dashed lines show the 95% confidence interval.

(Supplementary Information section 10, Supplementary Information Table 8).

Diffuse microseepage (24 Tg CH₄ yr⁻¹), macro-seeps and mud volcanoes (8.1 Tg CH₄ yr⁻¹), submarine seepage (>7 Tg CH₄ yr⁻¹) and geothermal manifestations (5.7 Tg CH₄ yr⁻¹) represent the main categories of natural geological CH₄ emissions in the latest comprehensive bottom-up analysis⁷. Each of these four categories is nearly equivalent to, or exceeds, our upper bound (at 95% confidence) on the total preindustrial-era geological CH₄ emissions (5.4 Tg CH₄ yr⁻¹). Emission estimates for diffuse microseepage are based on limited flux-chamber measurements in regions of known gas seepage (for example, ref. ²³), which are scaled up to a global flux estimate based on the total dryland area situated above hydrocarbon reservoirs (~10% of Earth's total land surface area), the percentage of measurements that show a positive flux, and emission rates chosen on the basis of several geological factors⁷. It is possible that the uncertainties associated with such global upscaling are much larger than reported, resulting in an overestimation by an order of magnitude or more. Similarly, emission estimates from macro-seeps, mud volcanoes and geothermal manifestations are derived from limited observations, which are scaled up to a global total⁷. To provide a sense of scale for the extrapolation in the case of mud volcanoes, ~0.0026 Tg CH₄ yr⁻¹ of measured CH₄ emissions (table S2 in the supplement of ref. ⁷) are scaled up to 6.1 Tg CH₄ yr⁻¹ (table 2 in ref. ⁷).

With regard to submarine seepage, recent studies suggest that CH₄ emissions to the atmosphere are probably very low owing to rapid dissolution of rising bubbles²⁴ and rapid oxidation of dissolved CH₄ (ref. ²⁵). Furthermore, ¹⁴CH₄ measurements in surface waters indicate minimal quantities of fossil CH₄ even in shallow waters over areas of active seeps or methane hydrate degradation²⁶. Our atmospheric ¹⁴CH₄ measurements for the preindustrial era indicate that either (1) the uncertainties associated with global upscaling of geological emissions from discrete measurements result in overestimation by an order of magnitude, or (2) geological CH₄ emissions quantified by these measurements were not present in the preindustrial era and may have been triggered by fossil fuel extraction from hydrocarbon reservoirs or other anthropogenic activity such as groundwater aquifer depletion. If the latter is true, such emissions cannot be considered natural.

A recent study³ used ice core ^δ¹³CH₄ measurements to arrive at a natural geological CH₄ emission estimate that is on par with what is indicated by bottom-up methods (~50 Tg CH₄ yr⁻¹). However, ref. ⁸ showed that ice core ^δ¹³CH₄ data do not provide a strong constraint on preindustrial-era geological emissions and are also compatible with a minimal geological source. Measurements of ethane²⁷ in ice cores have also been used to suggest considerable emissions of fossil CH₄ during the preindustrial era. However, this is also an ambiguous constraint because ice core measurements of ethane mole fraction cannot discriminate between contributions from biomass burning (a major source) and natural geological emissions¹¹. Our preindustrial-era ¹⁴CH₄ measurements, by contrast, place an unambiguous constraint on natural fossil CH₄ emissions by directly recording the ¹⁴C-free fraction of atmospheric CH₄.

Our ¹⁴CH₄ reconstruction does not allow accurate quantification of the post-1950 fossil CH₄ budget, owing to relatively poor constraints on the interfering nuclear ¹⁴CH₄ sources. Previous work used atmospheric ^δ¹³CH₄ measurements to quantify the global fossil CH₄ source in recent decades³, but relied on inventory-based assessments to constrain the natural geological component. We combine our ¹⁴CH₄-based constraint on natural geological emissions (1.6 Tg CH₄ yr⁻¹) with ^δ¹³CH₄-based constraints on the total fossil source (following the same one-box model approach as ref. ³; Supplementary Information section 11) to estimate recent anthropogenic fossil CH₄ emissions. This approach yields 177 ± 37 Tg CH₄ yr⁻¹ (1σ) for anthropogenic fossil CH₄ emissions during 2003–2012. Our estimate is 22% higher than the previous estimate of 145 ± 23 Tg CH₄ yr⁻¹ (1σ) over the same interval³, and 33–55% higher than the range of bottom-up estimates (114–133 Tg CH₄ yr⁻¹; ref. ²). We note that our ^δ¹³CH₄-based calculation uses an updated value for the CH₄ sink isotopic fractionation (Supplementary Information section 11); if we use the same value as ref. ³, the anthropogenic fossil source estimate is 194 ± 34 Tg CH₄ yr⁻¹ for the same time period.

Our results indicate that bottom-up inventories strongly underestimate CH₄ emissions from fossil fuel extraction, distribution and use. A study using both ground-based facility-scale measurements and verification from aircraft sampling found that US oil and natural-gas CH₄ emissions (largely from the production and gathering industry segments) are ~60% higher than those reported by the US Environmental Protection Agency²⁸, one of the primary data sources used in bottom-up inventories². If we consider a scenario in which the global bottom-up emissions of fossil CH₄ from the oil and natural-gas industries (79 Tg CH₄ yr⁻¹; ref. ²) are similarly underreported by 60%, this would amount to unreported emissions of ~47 Tg CH₄ yr⁻¹, which is in agreement with the fossil CH₄ emission shortfall that we identify in the current generation of bottom-up inventories (44–63 Tg CH₄ yr⁻¹). Our results imply that anthropogenic fossil CH₄ emissions now account for about 30% of the global CH₄ source and for nearly half of anthropogenic emissions, highlighting the critical role of emission reductions in mitigating climate change^{9,10}.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-1991-8>.

- Meinshausen, M. et al. Historical greenhouse gas concentrations for climate modelling (CMIP6). *Geosci. Model Dev.* **10**, 2057–2116 (2017).
- Saunio, M. et al. The global methane budget 2000–2012. *Earth Syst. Sci. Data* **8**, 697–751 (2016).
- Schwietzke, S. et al. Upward revision of global fossil fuel methane emissions based on isotope database. *Nature* **538**, 88–91 (2016); corrigendum **543**, 452 (2017).
- Lassey, K. R., Etheridge, D. M., Lowe, D. C., Smith, A. M. & Ferretti, D. F. Centennial evolution of the atmospheric methane budget: what do the carbon isotopes tell us? *Atmos. Chem. Phys.* **7**, 2119–2139 (2007).
- Zazzeri, G., Acuña Yeomans, E. & Graven, H. D. Global and regional emissions of radiocarbon from nuclear power plants from 1972 to 2016. *Radiocarbon* **60**, 1067–1081 (2018).
- Etiopie, G. *Natural Gas Seepage: The Earth's Hydrocarbon Degassing* Vol. 1 (Springer International Publishing, 2015).
- Etiopie, G., Ciotoli, G., Schwietzke, S. & Schoell, M. Gridded maps of geological methane emissions and their isotopic signature. *Earth Syst. Sci. Data* **11**, 1–22 (2019).
- Petrenko, V. V. et al. Minimal geological methane emissions during the Younger Dryas–Preboreal abrupt warming event. *Nature* **548**, 443–446 (2017).
- Höglund-Isaksson, L. Global anthropogenic methane emissions 2005–2030: technical mitigation potentials and costs. *Atmos. Chem. Phys.* **12**, 9079–9096 (2012).
- Howarth, R. W. Methane emissions and climatic warming risk from hydraulic fracturing and shale gas development: implications for policy. *Eng. Emis. Con. Tech.* **3**, 45–54 (2015).
- Nicewonger, M. R., Aydin, M., Prather, M. J. & Saltzman, E. S. Large changes in biomass burning over the last millennium inferred from paleoatmospheric ethane in polar ice cores. *Proc. Natl Acad. Sci. USA* **115**, 12413–12418 (2018).
- Bock, M. et al. Glacial/interglacial wetland, biomass burning, and geologic methane emissions constrained by dual stable isotopic CH₄ ice core records. *Proc. Natl Acad. Sci. USA* **114**, E5778–E5786 (2017).
- Lassey, K. R., Lowe, D. C. & Smith, A. M. The atmospheric cycling of radiomethane and the “fossil fraction” of the methane source. *Atmos. Chem. Phys.* **7**, 2141–2149 (2007).
- Etiopie, G., Milkov, A. V. & Derbyshire, E. Did geologic emissions of methane play any role in Quaternary climate change? *Global Planet. Change* **61**, 79–88 (2008).
- Petrenko, V. V. et al. Measurements of ¹⁴C in ancient ice from Taylor Glacier, Antarctica constrain in situ cosmogenic ¹⁴CH₄ and ¹⁴CO production rates. *Geochim. Cosmochim. Acta* **177**, 62–77 (2016).
- Petrenko, V. V. et al. ¹⁴CH₄ measurements in Greenland Ice: investigating last glacial termination CH₄ sources. *Science* **324**, 506–508 (2009).
- Severinghaus, J. P. et al. Deep air convection in the firn at a zero-accumulation site, central Antarctica. *Earth Planet. Sci. Lett.* **293**, 359–367 (2010).
- Buizert, C. et al. Gas transport in firn: multiple-tracer characterisation and model intercomparison for NEEM, Northern Greenland. *Atmos. Chem. Phys.* **12**, 4259–4277 (2012).
- Rommelaere, V., Arnaud, L. & Barnola, J.-M. Reconstructing recent atmospheric trace gas concentrations from polar firn and bubbly ice data by inverse methods. *J. Geophys. Res. Atmos.* **102**, 30069–30083 (1997).
- Trudinger, C. et al. Reconstructing atmospheric histories from measurements of air composition in firn. *J. Geophys. Res. Atmos.* **107**, (2002).
- Smil, V. *Energy Transitions: Global and National Perspectives* (ABC-CLIO, 2016).
- Hua, Q., Barbetti, M. & Rakowski, A. Z. Atmospheric radiocarbon for the period 1950–2010. *Radiocarbon* **55**, 2059–2072 (2013).
- Etiopie, G. & Klusman, R. W. Microseepage in drylands: flux and implications in the global atmospheric source/sink budget of methane. *Global Planet. Change* **72**, 265–274 (2010).
- McGinnis, D. F., Greinert, J., Artemov, Y., Beaubien, S. & Wüest, A. Fate of rising methane bubbles in stratified waters: how much methane reaches the atmosphere? *J. Geophys. Res. Oceans* **111**, C09007 (2006).
- Leonte, M. et al. Rapid rates of aerobic methane oxidation at the feather edge of gas hydrate stability in the waters of Hudson Canyon, US Atlantic Margin. *Geochim. Cosmochim. Acta* **204**, 375–387 (2017).
- Sparrow, K. J. et al. Limited contribution of ancient methane to surface waters of the U.S. Beaufort Sea shelf. *Sci. Adv.* **4**, eaao4842 (2018).
- Nicewonger, M. R., Verhulst, K. R., Aydin, M. & Saltzman, E. S. Preindustrial atmospheric ethane levels inferred from polar ice cores: a constraint on the geologic sources of atmospheric ethane and methane. *Geophys. Res. Lett.* **43**, 214–221 (2016).
- Alvarez, R. A. et al. Assessment of methane emissions from the U.S. oil and gas supply chain. *Science* **361**, 186–188 (2018).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Data availability

The ice core and firn air $^{14}\text{CH}_4$ data presented in Fig. 1 are provided in Supplementary Information Tables 2, 6. Additional measurements not provided in Supplementary Information Tables 1–8 are available via the NSF Arctic Data Center at <https://doi.org/10.18739/A2599Z216>.

Code availability

The code for the firn air inverse model and atmospheric box model (MATLAB) is available from the corresponding author upon request.

Acknowledgements This work was supported by US NSF awards OPP-1203779 (V.V.P.) OPP-1203686, OPP-0230452, ANT-0839031 (J.P.S.) ARC-1204084, ARC-1702920 (C.B.), a Packard Fellowship for Science and Engineering (V.V.P.), the National Institute of Water and Atmospheric Research through the Greenhouse Gases, Emissions and Carbon Cycle Science Programme (T.B.) and the Australian Government for the Centre for Accelerator Science at ANSTO through the National Collaborative Research Infrastructure Strategy (A.M.S.). We thank J. McConnell and P. Vallelonga for the interpretation of the ice core CFA data; P. Neff and E. Steig for sharing the ice-thinning model code; L. Davidge, J. Edwards, M. Pacicco and A. Adolph for assistance with firn air and ice core sampling; M. Jayred, L. Albershardt, T. Kuhl, D. Kirkpatrick and the US Ice Drilling programme for ice-drilling support; K. Gorham, J. Jenkins,

D. Einerson, Polar Field Services and the 109th New York Air National Guard for logistical support; the Australian Antarctic Science Program for supporting the Law Dome drilling and firn air sampling and CSIRO GASLAB, in particular R. Langenfelds, for analysis of the firn air sample trace gas concentrations.

Author contributions B.H. and V.V.P. designed the study and conducted field logistical and scientific preparations; B.H., V.V.P., M.N.D., C.B., P.F.P., R.B., J.S. and X.F. collected samples at Summit; B.H. measured $[\text{CO}]$ and extracted CH_4 and CO from firn air and ice core samples; C.B. developed the firn modelling code; B.H. and M.N.D. developed the box-model calculations; Q.H. and B.Y. graphitized the ^{14}C samples; A.M.S. measured ^{14}C ; P.F.P. and I.V. measured $\delta^{13}\text{CO}$; S.E.M. measured $\delta^{13}\text{CH}_4$; C.H. measured $[\text{CH}_4]$ and halogenated trace gases under the supervision of R.F.W.; E.D. supervised the firn air trace gas measurements; J.P.S. measured $\delta\text{Xe/Kr}$, $\delta\text{Kr/N}_2$, $\delta\text{Xe/N}_2$ and $\delta\text{Ne/N}_2$ and collected Megadunes firn air samples; R.B. measured the $\delta^{15}\text{N}$ of N_2 , the $\delta^{18}\text{O}$ of O_2 , $\delta\text{O}_2/\text{N}_2$ and $\delta\text{Ar/N}_2$; D.E. collected and supervised the analyses of the Law Dome firn air samples; T.B. extracted CH_4 from Megadunes and Law Dome samples; B.H. and V.V.P. analysed the data and B.H. drafted the manuscript with contribution from all authors.

Competing interests The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-1991-8>.

Correspondence and requests for materials should be addressed to B.H.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

A claustrum in reptiles and its role in slow-wave sleep

<https://doi.org/10.1038/s41586-020-1993-6>

Received: 13 August 2019

Accepted: 12 December 2019

Published online: 12 February 2020

Hiroaki Norimoto^{1,6}, Lorenz A. Fenk^{1,6}, Hsing-Hsi Li¹, Maria Antonietta Tosches^{1,2}, Tatiana Gallego-Flores¹, David Hain^{1,3}, Sam Reiter^{1,4}, Riho Kobayashi^{1,5}, Angeles Macias¹, Anja Arends¹, Michaela Klinkmann¹ & Gilles Laurent^{1✉}

The mammalian claustrum, owing to its widespread connectivity with other forebrain structures, has been hypothesized to mediate functions that range from decision-making to consciousness¹. Here we report that a homologue of the claustrum, identified by single-cell transcriptomics and viral tracing of connectivity, also exists in a reptile—the Australian bearded dragon *Pogona vitticeps*. In *Pogona*, the claustrum underlies the generation of sharp waves during slow-wave sleep. The sharp waves, together with superimposed high-frequency ripples², propagate to the entire neighbouring pallial dorsal ventricular ridge (DVR). Unilateral or bilateral lesions of the claustrum suppress the production of sharp-wave ripples during slow-wave sleep in a unilateral or bilateral manner, respectively, but do not affect the regular and rapidly alternating sleep rhythm that is characteristic of sleep in this species³. The claustrum is thus not involved in the generation of the sleep rhythm itself. Tract tracing revealed that the reptilian claustrum projects widely to a variety of forebrain areas, including the cortex, and that it receives converging inputs from, among others, areas of the mid- and hindbrain that are known to be involved in wake–sleep control in mammals^{4–6}. Periodically modulating the concentration of serotonin in the claustrum, for example, caused a matching modulation of sharp-wave production there and in the neighbouring DVR. Using transcriptomic approaches, we also identified a claustrum in the turtle *Trachemys scripta*, a distant reptilian relative of lizards. The claustrum is therefore an ancient structure that was probably already present in the brain of the common vertebrate ancestor of reptiles and mammals. It may have an important role in the control of brain states owing to the ascending input it receives from the mid- and hindbrain, its widespread projections to the forebrain and its role in sharp-wave generation during slow-wave sleep.

Slow-wave sleep and rapid-eye-movement (REM) sleep are the two main macroscopic components of electrophysiological sleep in mammals and birds^{4–6}, although some mammals may lack REM sleep⁷. The recent finding of alternating slow-wave and REM sleep in a reptile, the Australian bearded dragon *Pogona vitticeps*³, suggests that these two modes of sleep may predate the diversification of amniotes 320 million years ago. Sleep in *Pogona* is particularly interesting because the sleep cycle of this reptile is very short (3 minutes or less at room temperature), and is divided equally into slow-wave sleep and REM sleep³.

The dominant electrophysiological feature of *Pogona* slow-wave sleep is energy in the δ frequency band (around 0–4 Hz), which is caused by the reliable occurrence of sharp waves. Sharp waves typically contain a high-frequency ripple, forming a sharp-wave ripple complex (SWR)². SWRs were recorded from the DVR⁸—the dominant non-cortical pallial domain of sauropsid brains^{8–10}. REM sleep, by contrast, is characterized by broad-band energy, measured in the β band (10–40 Hz) in the cortex and DVR³.

Origin of sharp waves during slow-wave sleep

SWRs occurred reliably in the DVR during slow-wave sleep, and slow-wave sleep alternated regularly with REM sleep (Fig. 1a–c, Extended Data Fig. 1), as reported previously³. High-frequency ripples (around 70–150 Hz) rode on each sharp wave and contained action potentials. Local field potentials (LFPs) were highly correlated across DVR recording sites (peak correlation 0.74 over 18 h of slow-wave sleep, mean over two animals), but sharp waves that were recorded in the anterior medial pole of the DVR (amDVR) preceded their counterparts in more posterior or more lateral regions by up to 200 ms depending on the spacing between recording sites (Fig. 1d, e, Extended Data Fig. 1g, h), suggesting SWR propagation.

We next recorded from thick anterior transverse, horizontal and parasagittal slices of DVR in artificial cerebrospinal fluid solution (ACSF) (Methods, Extended Data Fig. 2a–f). All configurations produced

¹Max Planck Institute for Brain Research, Frankfurt am Main, Germany. ²Department of Biological Sciences, Columbia University, New York, NY, USA. ³Department of Life Sciences, Goethe University, Frankfurt am Main, Germany. ⁴Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan. ⁵Department of Neuropharmacology, Graduate School of Pharmaceutical Sciences, Nagoya City University, Nagoya, Japan. ⁶These authors contributed equally: Hiroaki Norimoto, Lorenz A. Fenk. ✉e-mail: gilles.laurent@brain.mpg.de

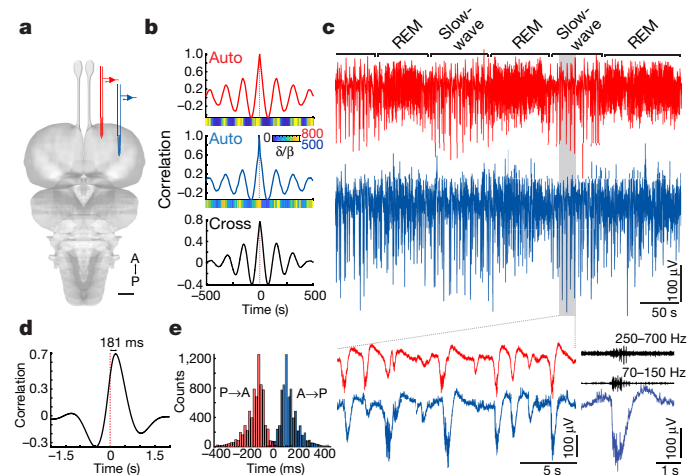


Fig. 1 | SWRs originate in the amDVR in sleeping *Pogona*. **a**, Simultaneous recordings from two sites in the DVR (subcortical). A, anterior; P, posterior. Scale bar, 1 mm. **b**, Auto- and cross-correlations of the δ/β power ratio as function of time from sites in **a**, calculated over 8 h of sleep. Coloured strips show the δ/β ratio over one single 1,000-s stretch of sleep. **c**, Short segment of the data that were analysed in **b** (same trace colours). Bottom left, magnification of a short segment of slow-wave sleep, illustrating SWR coordination and anterior–posterior delay. Bottom right, detail of a SWR and high-pass components (middle and top, black). **d**, Cross-correlation between broadband LFP waveforms (**c**) during 3.42 h of SWS. Reference trace is the anterior recording site: the anterior site leads. **e**, Delay distribution of sharp waves in the anterior (or posterior) DVR triggered on simultaneously recorded posterior (or anterior) DVR. See Methods and Extended Data Fig. 1.

spontaneous SWRs that matched those produced in sleep: a biphasic waveform (119 ± 40 ms) with a ripple (around 70–150 Hz) on the trough. SWRs in DVR slices were less frequent than those that occur during slow-wave sleep (12.4 ± 1.8 min⁻¹ (12 DVR slices, 10 animals) versus 16.45 ± 0.98 min⁻¹ during slow-wave sleep (5 slow-wave sleep epochs, 2 sleeping animals)), although not significantly so ($P = 0.18$, Student's *t*-test). SWR production in slices was not rhythmically interrupted by REM-sleep-like activity as it is during sleep. We patched 12 DVR neurons (Extended Data Fig. 2g–j) and found that, consistent with sleep data, they typically fired 0–3 action potentials during SWRs and were silent between SWRs. Under voltage clamp ($n = 2$), neurons displayed coincident excitatory and inhibitory input during sharp waves (with excitation dominating in current-clamp mode).

We also used multi-electrode arrays on DVR slices ($n = 3$ brains; Methods). As observed in vivo, SWRs propagated from anterior medial to posterior lateral poles (Fig. 2a–c). The apparent linear velocity of the wave in the slice plane was 39 mm s⁻¹, although propagation contained local angular components. We further divided DVR slices into ‘mini-slices’ ($n = 13$, Fig. 2d). Only those from the anterior medial pole produced SWRs, and rates did not differ significantly from controls (11.9 ± 1.7 min⁻¹; Fig. 2e, f).

scRNA-seq indicates a claustrum homologue

Using a single-cell RNA sequencing (scRNA-seq) strategy, we previously mapped the main neuronal types of the reptilian pallium¹¹ and described heterogeneity among glutamatergic cell types in the *Pogona* DVR. To characterize the amDVR, we sampled *Pogona* single cells more deeply and more extensively (Methods). Using unsupervised graph-based Louvain clustering on transcripts from 20,257 cells, we identified 4,054 pallial glutamatergic neurons that formed 29 glutamatergic clusters (Fig. 3a, Extended Data Fig. 3).

We located these clusters in the *Pogona* telencephalon using the expression of cluster-specific markers, which were detected by in situ

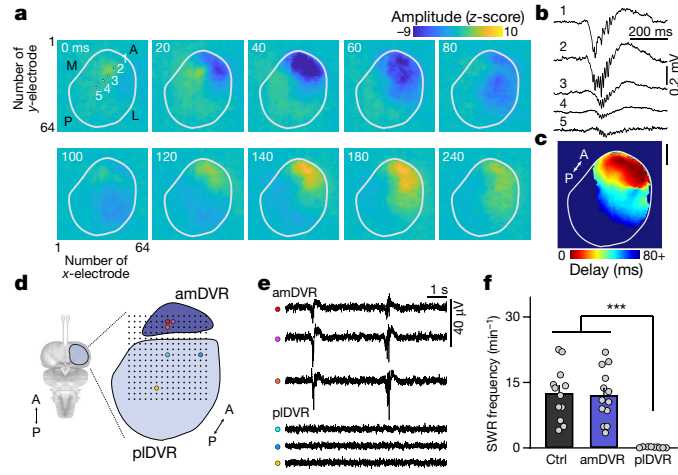


Fig. 2 | SWRs occur spontaneously in DVR slices and originate at the anterior medial pole. **a–c**, CMOS-MEA (complementary metal oxide semiconductor–microelectrode array) recordings of SWRs (see also Extended Data Fig. 2) propagating across a horizontal slice of the DVR (outlined). **a**, Instantaneous voltage samples at an interval of 20–60 ms. Squares 1–5 indicate the recording sites shown in **b**. Note the initiation at the anterior pole. L, lateral; M, medial. **b**, SWRs from sites 1–5 in **a**. Note the differences in amplitude and onset time between sites. **c**, Signal latency (relative to the earliest channel) over slice plane (mean of 12 SWRs; same slice as in **a**). Scale bar, 1 mm. **d–f**, SWRs in mini-slices. A 252-site microelectrode array, 200- μ m pitch. **d**, Thick horizontal slices of DVR were subdivided. **e**, Simultaneous LFPs recorded from coloured sites in **d**. pLDVR, posterior lateral DVR. **f**, Mean frequency of SWRs in intact slices (control (ctrl); $n = 12$ slices); amDVR ($n = 13$ mini-slices); and pLDVR ($n = 9$ mini-slices). *** $P < 0.001$. *P* values: control versus amDVR, $P = 1$, $t_{23} = 0.04$; control versus pLDVR, $P = 7.2 \times 10^{-6}$; $t_{19} = 6.3$; amDVR versus pLDVR, $P = 4.6 \times 10^{-6}$, $t_{19} = 6.3$ (two-sided Bonferroni test). Data are mean \pm s.e.m.

hybridization and/or immunohistochemistry¹¹. Two clusters (19 and 20, Fig. 3a) mapped to the amDVR, as shown by expression of *hpc4* (which encodes the calcium-binding protein hippocalcin) and *adarb2* (which encodes an RNA-editing enzyme), among others (Fig. 3b–d). Clusters 19 and 20 corresponded to medial and lateral subdivisions of the amDVR, as shown by expression of the copine-4 (*cpne4*) and nuclear hormone receptor (*rorb*) genes (Fig. 3e, f). When we repeated the mini-slice SWR recordings and labelled those slices afterwards with a hippocalcin antibody, we found that only hippocalcin-positive mini-slices from the anterior medial pole of the DVR generated SWRs (Extended Data Fig. 4).

Some amDVR markers (for example, *gng2*, *synpr* and *rgs12*; Fig. 3b) are known markers of the mammalian claustrum¹². To explore these molecular similarities further, we used Seurat v.3 to project *Pogona* single-cell transcriptomes onto mouse single-cell transcriptomes¹³ on the basis of a joint dimensionality reduction analysis¹⁴ (Methods). About 63% and 75% of amDVR cells (clusters 19 and 20, respectively) projected onto the mouse claustrum transcriptomic cluster (Fig. 3g). This suggests that—consistent with developmental observations^{10,15}—the *Pogona* amDVR and the mammalian claustrum are homologous.

To link our transcriptomic and physiological observations, we analysed the expression of genes that encode ion channels and neurotransmitter receptors in pallial glutamatergic clusters (143 genes detected in at least 20% of cells of at least one cluster; Methods). These genes were sufficient to distinguish the amDVR from other glutamatergic clusters (Extended Data Figs. 3, 5), and contained clusters of correlated genes (modules). One module with enriched expression in the amDVR (Fig. 3h) included receptors for noradrenaline, acetylcholine, dopamine and serotonin. In mammals, these neuromodulators influence sleep rhythms and are released by brain nuclei from the hypothalamus to the medulla^{4,5,16–18}. Glutamatergic neurons in the amDVR were among the

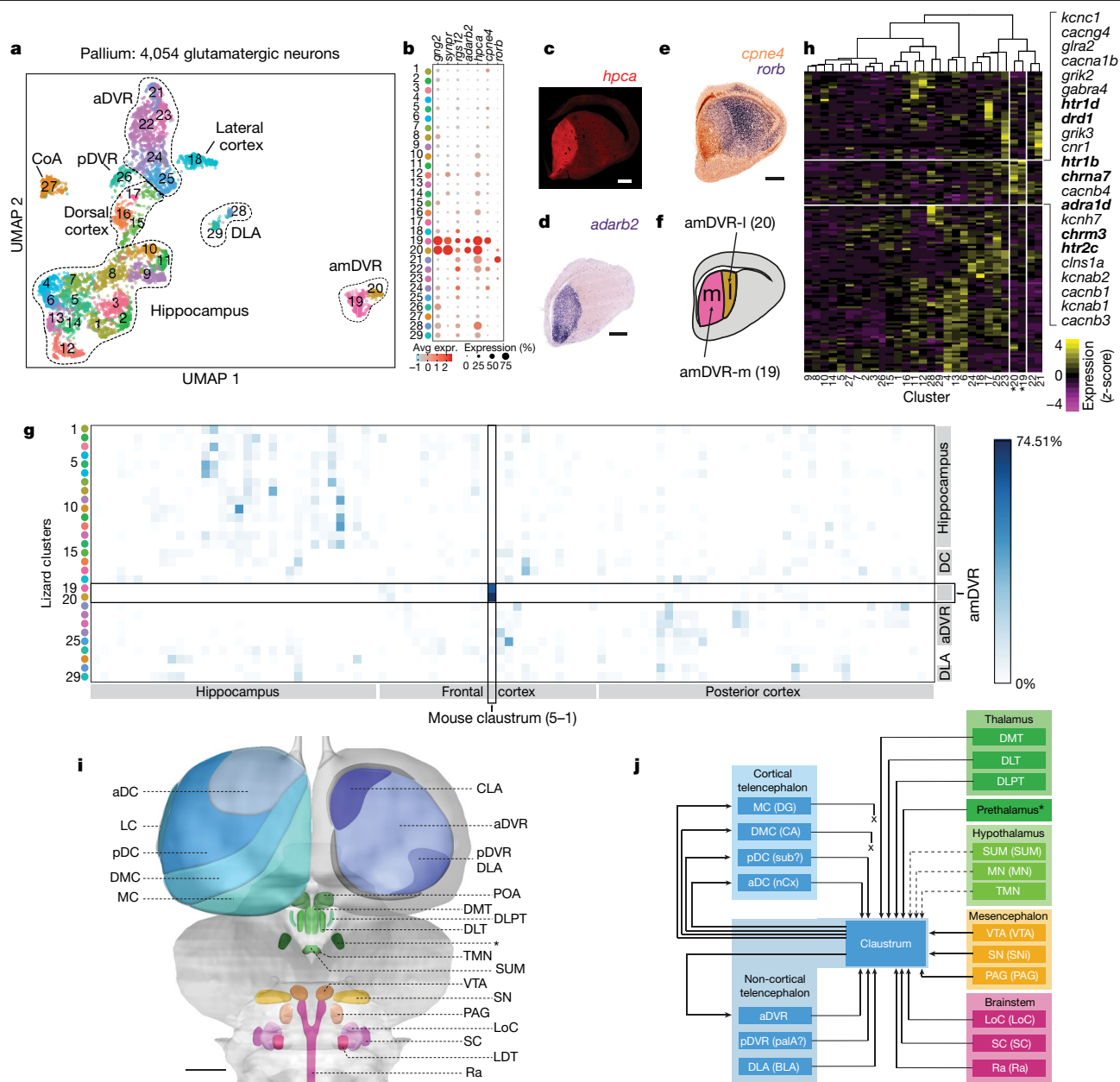


Fig. 3 | scRNA-seq and viral tract tracing show that the amDVR is a reptilian claustrum. **a**, Uniform manifold approximation and projection (UMAP)⁴² representation of single-cell transcriptomes of 4,054 *Pogona* glutamatergic pallial neurons. Cells are colour-coded by cluster (1–29). CoA, cortical amygdala; DLA, dorsal lateral amygdala. See Extended Data Figs. 3, 5. **b**, Expression across clusters of marker genes with high, specific expression in the amDVR (clusters 19, 20). These include markers of the mammalian claustrum. The size of the dot corresponds to the fraction of cells in which the gene is detected and the colour represents the expression level. **c–e**, Anterior transverse sections of *Pogona* telencephalon with immunostaining for hippocalin (*hpca*) (**c**), in situ hybridization for *adarb2* (**d**) and double in situ hybridization with *cpne4* and *rorb* probes (**e**). Scale bars, 500 μ m. **f**, Diagram showing the medial (m) and lateral (l) subdivisions of the amDVR. **g**, Transcriptomic similarity between lizard and mouse clusters, measured as the fraction of single-cell transcriptomes that map from *Pogona* to mouse clusters (Methods; mouse data are from a previous study¹³). **h**, Average expression in the *Pogona* clusters of 143 ion-channel and neurotransmitter-receptor genes (Extended Data Fig. 5). Clusters 19 and 20 (marked with an asterisk) correspond, respectively, to the medial and lateral subdivisions of the amDVR shown in f. Genes with enriched expression in the amDVR are listed on the right, with relevant neurotransmitter receptor genes in bold. **i**, Schematic

of the *Pogona* brain. Forebrain areas (blue) were identified by scRNA-seq and in situ hybridization¹¹. Diencephalic (green), mesencephalic (orange) and rhombencephalic (pink) areas were identified by immunohistochemistry and FISH (see Extended Data Fig. 6). aDC, anterior dorsal cortex; CLA, claustrum; DLT, dorso-lateral thalamus; DLPT, dorso-lateral posterior thalamus; DMC, dorso-medial cortex; DMT, dorso-medial thalamus; LC, lateral cortex; LDT, lateral dorsal tegmental nucleus; LoC, locus coeruleus; MC, medial cortex; PAG, periaqueductal grey; pDC, posterior dorsal cortex; POA, preoptic area; Ra, raphe nuclei; SC, subcoeruleus; SN, substantia nigra (SNi for mammals); SUM, supramammillary nucleus; TMN, tuberomammillary nucleus; VTA, ventral tegmental area. The prethalamus is marked with an asterisk, as in i. **j**, Summary of claustrum (amDVR) connectivity with areas in i, as determined by viral tracing. Arrows with solid lines indicate connections; 'x' indicates an absence of connection (absence of anterograde and retrograde labelling); and arrows with dotted lines indicate tentative connections (due to inconsistent labelling). Claustral projections to the pDVR or DLA were not conclusively tested, as we were not able to inject rAAV2-retro specifically into those small areas. Abbreviations for mammalian homologues are shown in parentheses: BLA, basolateral amygdala; CA, cornu ammonis; DG, dentate gyrus; nCx, neocortex; palA, pallial amygdala; sub, subiculum. The prethalamus is marked with an asterisk, as in i.

few types that co-expressed receptors for all four modulators (Extended Data Fig. 5). Hence, the amDVR expresses receptor types that are consistent with a sensitivity to input from circuits that control brain state.

The amDVR is extensively connected

We next mapped the connectivity of the amDVR with areas that have been associated with wake–sleep control in mammals (as suggested by the above data) and asked whether the amDVR connects widely with the rest of the pallium, as the claustrum does in mammals^{1,12,19–21}. We identified, where possible, the *Pogona* homologues of mammalian nuclei that are implicated in sleep^{4–6}. Relying on anatomical studies in related species (Methods), we used immunohistochemistry and fluorescence in situ hybridization (FISH) to identify and map these nuclei in the *Pogona* diencephalon, midbrain and brainstem (Fig. 3j, Extended Data Fig. 6), together with telencephalic areas mapped by scRNA-seq¹¹.

We mapped amDVR connectivity by local tracer injections²², using an adeno-associated virus vector (rAAV2-retro²³) carrying a fluorescent-protein gene under the CAG or hSyn promoter for (mostly) retrograde labelling (Methods). rAAV2-retro was sometimes co-injected with the (mostly) anterograde tracer AAV2/9-CB7-mCherry-WPRE for injection-site identification. Because they do not cross synapses^{23–25}, these tracers revealed the direct targets (AAV2/9-CB7) and sources (rAAV2-retro) of the injection site. The results are summarized in Fig. 3j. The names on the left all describe telencephalic structures, for which the input and output connectivity with the amDVR ('claustrum') could be tested. On the right are deeper structures that for anatomical reasons could not be reached for injection. For these structures, connectivity to the claustrum was established only by retrograde labelling from the amDVR, and the question of whether the claustrum projects to those areas will require further investigation and direct demonstration.

The cortical sources of input to the amDVR were the anterior and posterior dorsal cortices (Fig. 3j, Extended Data Fig. 6c). Retrograde and anterograde tracers revealed no direct projections from the dorso-medial cortex (homologue of the hippocampal cornu ammonis (CA)) and medial cortex (homologue of the dentate gyrus (DG)) to the amDVR, even though the amDVR projects to both (Fig. 3j). In the subcortical pallium, the anterior DVR (aDVR) and posterior DVR (pDVR) showed strong projections to the amDVR. The amDVR also received input from the dorsal thalamus (dorso-medial, dorso-lateral and dorso-lateral posterior nuclei), prethalamus, hypothalamus, ventral tegmental area, substantia nigra, the periaqueductal grey in the midbrain, and the locus coeruleus, subcoeruleus and raphe nucleus in the brainstem (Extended Data Fig. 6).

The amDVR projected to the hippocampus (medial cortex and dorso-medial cortex), posterior dorsal cortex (potential subiculum homologue) and anterior dorsal cortex (neocortex homologue)¹¹. In the subcortical pallium, projections to the aDVR were dense and extensive, consistent with sharp-wave propagation (Figs. 1, 2). Projections between the amDVR and some of its targets appeared ordered: more lateral amDVR projected to rostral aDVR and central amDVR projected to caudal aDVR. Conversely, input to the amDVR from the cortex (anterior and posterior dorsal cortex) was strongest laterally and weakest medially (absent from dorso-medial and medial cortices, or hippocampus).

Hence, the amDVR is connected with the pallial forebrain and receives input from areas that are implicated in wake–sleep control—consistent with the widespread expression of many receptor genes that are specific to these areas. On the basis of these transcriptomic and anatomical data, we conclude that the amDVR is the reptilian homologue of the mammalian claustrum.

The claustrum homologue in turtles

Having applied similar transcriptomic approaches to those used in *Pogona* to the turtle *T. scripta*¹¹—a species on a distant branch of the

reptilian tree—we looked for a turtle claustrum. Comparison of transcriptomic data (Methods) yielded four potential clusters in *Trachemys* (Extended Data Fig. 7). Cells in these clusters were located in a region known as the pallial thickening^{11,26,27}. Turtle pallial thickening and lizard amDVR are both in the anterior pallium, consistent with their similar developmental origin in anterior lateral pallium¹⁰; however, turtle pallial thickening is lateral to the aDVR and close to the olfactory cortex, rather than being fused to the rest of the DVR as the claustrum is in *Pogona*. Architectonics also differed: the *Pogona* claustrum is nuclear and composed of isotropically distributed multipolar neurons, whereas turtle pallial thickening forms a curved sheet that extends the anterior dorsal cortex and is traversed from below by lateral geniculate nucleus (LGN) axons en route to the visual cortex²⁷. Principal neurons in turtle pallial thickening (revealed after rAAV2-retro injection into the dorso-medial cortex) are pyramid-like, with apical and basal dendrites (Extended Data Fig. 7d). Despite these differences, slices of turtle pallial thickening produced SWRs that led those in the DVR, as in *Pogona*. This pallial thickening therefore appears to be the turtle claustrum, suggesting that a homologue of the claustrum already existed in the common ancestor of amniotes.

Manipulating claustrum activity

We developed a reduced ex vivo preparation of the *Pogona* forebrain, which enabled direct access to the non-cortical pallium after removal of the cortex (Methods). This preparation generated spontaneous SWRs in the claustrum and DVR that were similar to those recorded in vivo during sleep and to those that occur in slices containing both claustrum and DVR (claustrum + DVR) (Extended Data Fig. 8). SWRs occurred continuously but more frequently in the forebrain preparation ($21.6 \pm 5.4 \text{ min}^{-1}$, 4 brains) than in slices ($12.4 \pm 1.8 \text{ min}^{-1}$, $n = 13$). SWRs in the claustrum led those in the DVR (Extended Data Fig. 8f), with delays similar to those observed during sleep or in slices of claustrum + DVR (11–141 ms, peak mean correlation 0.57, 4 brains). To test the causal role of the claustrum in generating SWRs, we injected tetrodotoxin (TTX) selectively into the claustrum ex vivo ($n = 4$, 3 animals). This resulted in a prolonged silencing of the claustrum, and the concomitant cessation of SWRs in the ipsilateral DVR (Extended Data Fig. 8b–d).

We next generated lesions in one or both claustra in vivo using ibotenic acid (Methods; three animals). Bilateral recordings from the DVR in sleeping lesioned animals revealed that the rhythmic modulation of β -band activity (REM sleep) was unaffected, but that SWRs (characteristic of slow-wave sleep) were eliminated on the side(s) of the lesioned claustrum (Fig. 4a–d, Extended Data Fig. 9). These findings show that the claustrum is required for the production of SWRs in the DVR during slow-wave sleep; that its action is unilateral; and that it is not involved in the alternating sleep rhythm itself.

Because the claustrum receives direct input from areas that are implicated in wake–sleep control in mammals and expresses receptors for their transmitters (Fig. 3j, Extended Data Fig. 5), we tested how sensitive SWR production was to those transmitters^{4–6,16}. Dopamine significantly increased the rate of SWR production, whereas acetylcholine and serotonin decreased it (Fig. 4e). We selected serotonin for further experiments. Consistent with tracing data that indicated a serotonergic input from the raphe nuclei, the claustrum contained serotonin-positive fibres (Extended Data Fig. 10a). Serotonin at concentrations of $1 \mu\text{M}$ or higher suppressed SWRs ($n = 9$ claustrum + DVR slices, 9 animals; Extended Data Fig. 10b). This effect was best mimicked by the serotonin receptor-1D agonist L703,664 (Fig. 4f), consistent with scRNA-seq results (Extended Data Fig. 5). We then superfused slices with caged serotonin (Methods). SWRs were suppressed within seconds of the onset of illumination, and resumed when illumination stopped (Fig. 4g, h).

The mammalian claustrum is hypothesized to have a role in higher cognition^{1,28,29} because of its hub-like connectivity^{12,30–32}. Direct

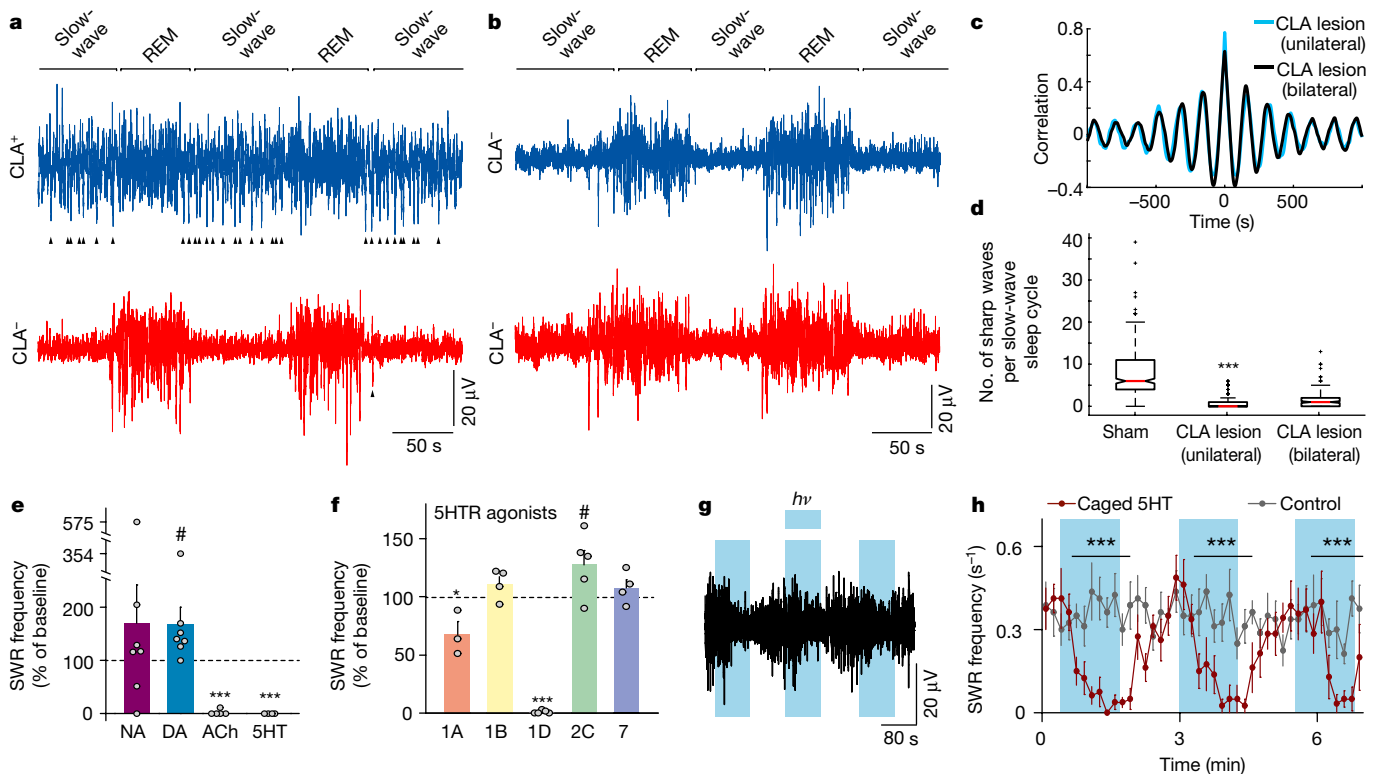


Fig. 4 | SWR production in the DVR depends on claustrum integrity and modulation. **a–d**, Ibotenic-acid-induced lesions of the claustrum and SWRs in sleeping lizards (see also Extended Data Fig. 9). **a**, Short sleep segment showing LFP (<150 Hz) from left and right DVRs after unilateral claustrum lesion. Sham-lesioned hemisphere (CLA⁺) is shown in blue; lesioned hemisphere (CLA⁻) is shown in red. The arrowheads indicate sharp waves in the DVR. Sleep rhythm is intact but sharp waves are nearly absent in slow-wave sleep on the lesioned side. **b**, Same as **a**, in a lizard with bilateral lesions of the claustrum. Note the absence of sharp waves. **c**, Cross-correlation of β -band (REM sleep) power across hemispheres in lesioned animals. **d**, Number of sharp waves per slow-wave sleep cycle in sham and claustrum-lesioned hemispheres. ***, significantly different from sham; $P < 1.73 \times 10^{-60}$, $W = 64,252$ (Wilcoxon signed-rank test; data from 2 animals, 4 nights, 375 cycles). For details of box plots, see Methods. **e–h**, Experiments in slices of DVR + claustrum or isolated claustrum. **e**, Effects

of superfused noradrenaline (NA; 25 μ M; $n = 7$), dopamine (DA) agonist SKF38393 (10 μ M; $n = 7$), acetylcholine (ACh) agonist carbachol (50 μ M; $n = 5$) and serotonin (5HT; 10 μ M; $n = 4$) on the frequency of spontaneous SWRs. **f**, Action of serotonin receptor (5HT₁) agonists on the rate of spontaneous SWRs in isolated slices of the claustrum. $n = 3$ experiments (5HT_{1A}); $n = 4$ (5HT_{1B}); $n = 5$ (5HT_{1D}); $n = 5$ (5HT_{2C}); $n = 4$ (5HT₇). ***, * and #, significantly different from baseline; *** $P = 8.0 \times 10^{-3}$, $T = 15$ (two-sided Wilcoxon rank-sum test); * $P = 0.04$, $t_4 = -2.9$; # $P = 0.049$, $t_8 = -2.3$ (paired t -test). Data are mean \pm s.e.m. (**e**, **f**). **g**, Light-triggered uncaging of serotonin suppresses spontaneous SWRs in CLA + DVR slices. Blue shading, light on ($h\nu$). **h**, Summary of eight experiments performed as in **g**, with bins of 10 s. Data are mean \pm s.e.m. For the control experiments, light pulses were applied to ACSF-superfused slices. ***, experimental bins significantly different from control; $P = 1.5 \times 10^{-4}$, $T = 36$ (two-sided Mann–Whitney rank-sum test).

experimental tests are difficult, however, owing to the anatomy of the claustrum^{12,33}. Using scRNA-seq and tract-tracing techniques, we identified a claustrum in two reptiles from distant evolutionary lineages. This result, added to mammalian evidence, suggests that a claustrum probably existed in the common ancestor of amniotes. The claustrum probably derives from the lateral pallium and may correspond to parts of the mesopallium in birds^{34,35}. Thus, if the claustrum has a role in higher cognition in mammals, this role may be derived from other functions in a common amniote ancestor. The claustrum assumes different architectonics, which are reflected in neuronal morphology, in two distant reptiles. (Of note, differences also exist between marsupial and placental mammals³⁶.) Because the claustrum produces SWRs in both reptiles, architectonics probably have little role in SWR generation.

The claustrum participates in the generation and relaying of SWRs that are characteristic of slow-wave sleep in *Pogona*. Given the widespread connectivity of the claustrum and its input from wake–sleep-controlling areas, it may be implicated in coordinating forebrain states during sleep. Early experiments in cats³⁷ described sleep-like behaviour after (though not during) low-frequency stimulation of the claustrum. These results remain uncertain because selective stimulation of the mammalian claustrum is difficult. More recent results in rodents, using markers of synaptic activity³⁸, suggest that the claustrum is active

during REM sleep. Other studies^{39,40} suggest that the claustrum acts to shut down the cortex, through dominant projections onto cortical interneurons. This action would cause a general cortical down-state, as is possibly seen during certain phases of slow-wave sleep⁴⁰. These results collectively suggest a tentative link between the claustrum and sleep in mammals.

During sleep in *Pogona*, SWRs originate in the claustrum and propagate to the rest of the non-cortical pallium—the mammalian homologue of the amygdaloid complex¹¹. By virtue of ascending input from areas that control the wake–sleep cycle, the claustrum is ideally positioned to act as a relay for wake–sleep-related states in the forebrain. During sleep the claustrum alternates between SWR production and REM, presumably driven by alternating ascending inputs that are themselves independent of claustrum integrity. Claustrum projections suggest a distributed action on the cortex, hippocampus, amygdala and other areas of the forebrain. SWRs in sleeping *Pogona* in vivo are each correlated with a short phasic inhibition of the cortex (consistent with stimulation experiments (Extended Data Fig. 8) and with results in rodents^{39,40}) followed by cortical excitation³ (consistent with coordination between area CA1 and the medial prefrontal cortex in rodents⁴¹). The mechanisms that underlie this coordination, and the nature of sleep-related inputs to the claustrum, now require characterization.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-1993-6>.

- Crick, F. C. & Koch, C. What is the function of the claustrum? *Phil. Trans. R. Soc. Lond. B* **360**, 1271–1279 (2005).
- Buzsáki, G. Hippocampal sharp wave-ripple: a cognitive biomarker for episodic memory and planning. *Hippocampus* **25**, 1073–1188 (2015).
- Shein-Idelson, M., Ondracek, J. M., Liaw, H. P., Reiter, S. & Laurent, G. Slow waves, sharp waves, ripples, and REM in sleeping dragons. *Science* **352**, 590–595 (2016).
- Saper, C. B. & Fuller, P. M. Wake-sleep circuitry: an overview. *Curr. Opin. Neurobiol.* **44**, 186–192 (2017).
- Weber, F. & Dan, Y. Circuit-based interrogation of sleep control. *Nature* **538**, 51–59 (2016).
- Scammell, T. E., Arrigoni, E. & Lipton, J. O. Neural circuitry of wakefulness and sleep. *Neuron* **93**, 747–765 (2017).
- Lyamin, O. I., Manger, P. R., Ridgway, S. H., Mukhametov, L. M. & Siegel, J. M. Cetacean sleep: an unusual form of mammalian sleep. *Neurosci. Biobehav. Rev.* **32**, 1451–1484 (2008).
- Naumann, R. K. & Laurent, G. in *Evolution of Nervous Systems* Vol. 1 (ed. Kaas, J. H.) 491–518 (Elsevier, 2017).
- Moreno, N. & González, A. Evolution of the amygdaloid complex in vertebrates, with special reference to the anamnio-amniotic transition. *J. Anat.* **211**, 151–163 (2007).
- Puelles, L. et al. in *Evolution of Nervous Systems* Vol. 1 (ed. Kaas, J. H.) 519–555 (Elsevier, 2017).
- Tosches, M. A. et al. Evolution of pallium, hippocampus, and cortical cell types revealed by single-cell transcriptomics in reptiles. *Science* **360**, 881–888 (2018).
- Wang, Q. et al. Organization of the connections between claustrum and cortex in the mouse. *J. Comp. Neurol.* **525**, 1317–1346 (2017).
- Saunders, A. et al. Molecular diversity and specializations among the cells of the adult mouse brain. *Cell* **174**, 1015–1030 (2018).
- Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 (2019).
- Striedter, G. F. The telencephalon of tetrapods in evolution. *Brain Behav. Evol.* **49**, 179–213 (1997).
- Monti, J. M. Serotonin control of sleep-wake behavior. *Sleep Med. Rev.* **15**, 269–281 (2011).
- Clément, O., Sapin, E., Béréd, A., Fort, P. & Luppi, P. H. Evidence that neurons of the sublaterodorsal tegmental nucleus triggering paradoxical (REM) sleep are glutamatergic. *Sleep* **34**, 419–423 (2011).
- Hobson, J. A., McCarley, R. W. & Wyzinski, P. W. Sleep cycle oscillation: reciprocal discharge by two brainstem neuronal groups. *Science* **189**, 55–58 (1975).
- da Costa, N. M., Fürsinger, D. & Martin, K. A. The synaptic organization of the claustral projection to the cat's visual cortex. *J. Neurosci.* **30**, 13166–13170 (2010).
- Druga, R. in *The Claustrum: Structural, Functional, and Clinical Neuroscience* (eds Smythies, J. R. et al.) 29–84 (Academic, 2014).
- Olson, C. R. & Graybiel, A. M. Sensory maps in the claustrum of the cat. *Nature* **288**, 479–481 (1980).
- Pammer, L. *Explorations of Turtle Cortex Function through Molecular, Optogenetic and Electrophysiological Techniques*. PhD thesis, Goethe Univ. (2017).
- Tervo, D. G. et al. A designer AAV variant permits efficient retrograde access to projection neurons. *Neuron* **92**, 372–382 (2016).
- Oh, S. W. et al. A mesoscale connectome of the mouse brain. *Nature* **508**, 207–214 (2014).
- Harris, J. A., Oh, S. W. & Zeng, H. Adeno-associated viral vectors for anterograde axonal tracing with fluorescent proteins in nontransgenic and Cre driver mice. *Curr. Protoc. Neurosci.* **59**, 1.20.1–1.20.18 (2012).
- Desan, P. H. in *The Forebrain of Reptiles* (eds Schwedtfeger, W. K. & Smeets, W. J.) 1–11 (Karger, 1987).
- Heller, S. B. & Ulinski, P. S. Morphology of geniculocortical axons in turtles of the genera *Pseudemys* and *Chrysemys*. *Anat. Embryol.* **175**, 505–515 (1987).
- Atlan, G. et al. The claustrum supports resilience to distraction. *Curr. Biol.* **28**, 2752–2762 (2018).
- Smythies, J., Edelstein, L. & Ramachandran, V. Hypotheses relating to the function of the claustrum. *Front. Integr. Neurosci.* **6**, 53 (2012).
- Dillingham, C. M., Janowski, M. M., Chandra, R., Frost, B. E. & O'Mara, S. M. The claustrum: considerations regarding its anatomy, functions and a programme for research. *Brain Neurosci. Adv.* **1**, 1–9 (2017).
- Edelstein, L. R. & Denaro, F. J. The claustrum: a historical review of its anatomy, physiology, cytochemistry and functional significance. *Cell. Mol. Biol.* **50**, 675–702 (2004).
- Goll, Y., Atlan, G. & Citri, A. Attention: the claustrum. *Trends Neurosci.* **38**, 486–495 (2015).
- Mathur, B. N., Caprioli, R. M. & Deutch, A. Y. Proteomic analysis illuminates a novel structural definition of the claustrum and insula. *Cereb. Cortex* **19**, 2372–2379 (2009).
- Puelles, L. in *The Claustrum: Structural, Functional, and Clinical Neuroscience* (eds Smythies, J. R. et al.) 119–176 (Academic, 2014).
- Briscoe, S. D., Albertin, C. B., Rowell, J. J. & Ragsdale, C. W. Neocortical association cell types in the forebrain of birds and alligators. *Curr. Biol.* **28**, 686–696 (2018).
- Buchanan, K. J. & Johnson, J. I. Diversity of spatial relationships of the claustrum and insula in branches of the mammalian radiation. *Ann. NY Acad. Sci.* **1225**, E30–E63 (2011).
- Gabor, A. J. & Peele, T. L. Alterations of behavior following stimulation of the claustrum of the cat. *Electroencephalogr. Clin. Neurophysiol.* **17**, 513–519 (1964).
- Renouard, L. et al. The supramammillary nucleus and the claustrum activate the cortex during REM sleep. *Sci. Adv.* **1**, e1400177 (2015).
- Jackson, J., Karnani, M. M., Zemelman, B. V., Burdakov, D. & Lee, A. K. Inhibitory control of prefrontal cortex by the claustrum. *Neuron* **99**, 1029–1039 (2018).
- Narikiyo, K. et al. The claustrum coordinates cortical slow-wave activity. Preprint at *bioRxiv* <https://doi.org/10.1101/286773> (2018).
- Siapas, A. G. & Wilson, M. A. Coordinated interactions between hippocampal ripples and cortical spindles during slow-wave sleep. *Neuron* **21**, 1123–1128 (1998).
- McInnes, L., Healy, J. & Melville, J. UMAP: Uniform manifold approximation and projection for dimension reduction. Preprint at <http://arxiv.org/abs/1802.03426> (2018).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Methods

Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Animals

Lizards (*Pogona vitticeps*) of either sex, weighing 100–400 g, were obtained from our institute colony, selected for sex, size, weight, health status and wild-type colouring. Wild-type turtles (*T. scripta elegans* or *Chrysemys picta*) of either sex, weighing 200–400 g, were obtained from an open-air breeding colony (NASCO Biology). The animals were housed in our state-of-the-art animal facility.

All experimental procedures were performed in accordance with German animal welfare guidelines: permit no. V54-19c 20/15-F126/1005 delivered by the Regierungspräsidium Darmstadt (E. Simon).

Lizard surgery for chronic recordings

Twenty-four hours before surgery, the lizard was administered analgesics (butorphanol, 0.5 mg kg⁻¹ subcutaneously; meloxicam, 0.2 mg kg⁻¹ subcutaneously) and antibiotics (marbofloxacin, marbocyl, 2 mg kg⁻¹). On the day of surgery, anaesthesia was initiated with isoflurane, and maintained with isoflurane (1–4 vol%) after intubation. The lizard was placed in a stereotactic apparatus after ensuring deep anaesthesia (absence of corneal reflex). Body temperature during surgery was maintained at 32 °C using a heating pad and oesophageal temperature probe. Heart rate was monitored using a Doppler flow detector. The skin covering the skull was disinfected using 10% povidone-iodine solution before removal with a scalpel. A small (around 3 × 2-mm) craniotomy was then drilled posterior lateral to the parietal eye along the midline. The dura and arachnoid layers covering the forebrain were removed with fine forceps, and the pia was removed gently over the area of electrode insertion (dorsal or dorso-medial cortex). The exposed skull was covered with a layer of ultraviolet (UV)-hardening glue, and the bare ends of two insulated stainless steel wires were secured in place subdurally with UV-hardening glue to serve as the reference and ground.

For insertion of silicon probes, probes were mounted on a Nanodrive (Cambridge Neurotech) and secured to a stereotactic adaptor. On the day after the surgery, probes were slowly lowered into the tissue (about 0.9–1.2 mm). The brain was covered with Duragel followed by Vaseline. After connecting grounds, the skull, craniotomy and probes were secured with dental cement. After surgery, lizards were released from the stereotactic apparatus and left on a heating pad set to 32 °C until full recovery from anaesthesia.

In vivo electrophysiology

One week before surgery, lizards were habituated to a sleep arena for a minimum of two nights. One to two hours before lights off, the lizard was placed in the sleep arena, which was itself placed in a 3 × 3 × 3-m EM-shielded room. The animal was left to sleep and behave naturally overnight, and returned to its home terrarium 3–4 h after lights on. The animal then received food and water. Recordings were made from the cortex, anterior DVR (including claustrum) and/or posterior DVR of chronically implanted adult lizards. Electrodes were 32-channel silicon probes (50-µm pitch, 177-µm² surface area for each site; in 2 rows of 16 contacts).

Recordings were performed with a Cheetah Digital Lynx SX system and HS-36 headstages of unity gain and high input impedance (~1 TΩ). The headstage was connected with a headstage adaptor to a connector on the head, and a lightweight shielded tether cable connected the headstage to the acquisition system. Recordings were grounded and referenced against one of the reference wires. Signals were sampled at 32 kHz, with wide-band 0.1–9,000 Hz. Electrophysiological traces were typically low-pass filtered at 150 Hz with a 2-pole Butterworth filter for display.

Ibotenic acid lesion experiments

In preparation for claustrum lesion experiments we carefully removed, using fine forceps in anaesthetized animals, the pia overlaying the dorsal cortex and inserted a beveled quartz micropipette at an angle of 90° to the surface, to a depth of 1,050–1,150 µm from the surface, at appropriate anterior–posterior and medial–lateral coordinates to reach the centre of the claustrum. Ibotenic acid (400–600 nM; 5 µg µl⁻¹ in PBS, pH 7.2) was injected at a rate of 50–100 nl min⁻¹ (UMP3, World Precision Instruments). The injection pipette was retracted 3 min after the end of injection. Two silicon recording probes were subsequently positioned bilaterally, as described above, for DVR recordings. For sham claustrum lesions, we injected PBS alone (same methods and volumes) on the sham-lesion side. Recordings were carried out each night from one to six days after surgery. Effects of the lesions could already be observed 24 h after surgery. A week after each experiment, the animal was killed and its brain was sectioned and stained with Nissl for histological confirmation.

SWR delay calculation

Sharp waves were detected as described previously (template-based detection³). After independently detecting SWRs on probes in the aDVR and pDVR throughout the dataset, the delay between SWRs across probes was calculated by pairing SWRs on one probe with the SWR closest in time on the second probe. Pairs occurring more than 500 ms apart were ignored.

SWRs at the slow-wave sleep–REM sleep transition point

REM and slow-wave sleep periods and the timing of their transition were calculated as described previously³. Average SWR rates and amplitudes were calculated by averaging the values triggered on all slow-wave sleep–REM sleep transition points within 100-ms bins, and smoothing the resulting histogram with a Gaussian filter (s.d., 25 ms).

In ibotenic acid lesion experiments, sleep cycles were determined using median filtered β-band power (10–40 Hz, as above), for a 6-h period beginning 3 h after the recording start time. The time course of β was filtered above 0.001 Hz with a 2-pole Butterworth filter, and additionally smoothed with a Gaussian filter (s.d., 20 s). Periods of slow-wave sleep were conservatively defined as ones in which this signal was less than 1 s.d. below the mean. To avoid false sharp-wave detections in lesioned animals (which demonstrate reduced low-frequency power), sharp waves were detected by thresholding the voltage trace (1.5–2.5 s.d. below the mean) after low-pass filtering at 4 Hz with a 2-pole Butterworth filter. The threshold was adapted to each lesion experiment and was the same for both hemispheres within each experiment.

Sharp-wave shape statistics

For comparison with ex vivo and slice sharp waves, sharp waves detected in vivo were low-pass filtered at 20 Hz using a 2-pole Butterworth filter.

Ex vivo brain and slice preparations

Adult lizards or turtles were deeply anaesthetized with isoflurane, ketamine (60 mg kg⁻¹) and midazolam (2 mg kg⁻¹). After loss of the corneal reflex, the animals were decapitated and the heads were rapidly transferred into cooled ACSF solution (lizard: 126 mM NaCl, 3 mM KCl, 1.8 mM CaCl₂, 4 mM MgCl₂, 24 mM NaHCO₃, 0.72 mM NaH₂PO₄, 20 mM glucose, pH 7.4; turtle: 96.5 mM NaCl, 2.6 mM KCl, 4 mM CaCl₂, 2 mM MgCl₂, 31.5 mM NaHCO₃, 20 mM glucose, pH 7.4) bubbled with carbogen gas (95% O₂, 5% CO₂).

Ex vivo intact subcortical slabs were prepared with iridectomy scissors, after isolation of the lizard brain. For slice preparation, coronal, horizontal or sagittal subcortical area slices (700 µm thick) were cut using a vibratome (VT 1200S, Leica) in ice-cold oxygenated ACSF. The slices were allowed to recover for at least 60 min and then submerged in a chamber filled with oxygenated ACSF (lizards: 126 mM NaCl, 3 mM

Article

KCl, 1.8 mM CaCl_2 , 1 mM MgCl_2 , 24 mM NaHCO_3 , 0.72 mM NaH_2PO_4 , 20 mM glucose, pH 7.4; turtle: 96.5 mM NaCl, 2.6 mM KCl, 4 mM CaCl_2 , 2 mM MgCl_2 , 31.5 mM NaHCO_3 , 20 mM glucose, pH 7.4) at 20–22 °C.

Ex vivo brain and slice physiology and SWR detection

During recordings, oxygenated ACSF (lizard: 126 mM NaCl, 3 mM KCl, 1.8 mM CaCl_2 , 1.2 mM MgCl_2 , 24 mM NaHCO_3 , 0.72 mM NaH_2PO_4 , 20 mM glucose, pH 7.4; turtle: 96.5 mM NaCl, 2.6 mM KCl, 4 mM CaCl_2 , 2 mM MgCl_2 , 31.5 mM NaHCO_3 , 20 mM glucose, pH 7.4) was constantly superfused at 18–20 °C (ex vivo) and 18–21 °C (slices) at 4 ml min⁻¹. LFPs were recorded using microelectrode arrays, silicon probes or glass pipettes filled with ACSF. The electrodes were carefully placed in the targeted areas with micromanipulators. Signals were low-pass filtered at 2 kHz and digitized at 20 kHz. For analysis of sharp waves, the traces were further low-pass filtered at 20 Hz using a 2-pole Butterworth filter. SWRs were detected at a threshold of 3× s.d. of the total signal. The detected events were visually scrutinized and manually rejected if they were erroneously detected. Events lasting less than 30 ms were also discarded as they were typically artefacts. For claustrum electrical-stimulation experiments, stimulation pulses lasted 50 µs and were delivered with bipolar electrodes. Multi-unit extracellular recordings in cortex were carried out with glass micropipettes filled with ACSF. Mini-slices were cut with a sharp razor blade and were 0.61–3.12 mm² in surface area.

CMOS–MEA experiments

The slices were placed over a high-density microelectrode array (3Brain AG) of 4,096 electrodes (electrode size, 21 × 21 µm; pitch, 81 µm; 64 × 64 matrix; 5.12 × 5.12-mm area). During recording, ACSF perfusion was interrupted to avoid movements of the slices and noise as a result of ACSF flux. Signals were sampled at 18 kHz with a high-pass filter at 1 Hz.

Saturating or damaged channels were detected as channels whose voltage crossed ± 500 µV and were removed from later analysis. Channel data were low-pass filtered at 20 Hz and z-scored, and troughs greater than 5(z) below the mean on the channel with the largest signal were taken as sharp waves. The signal ± 400 ms from these peak times, on all channels, was taken as a SWR episode. For calculation of SWR latency, SWRs were averaged on each channel and the time that the average signal crossed 1(z) below the mean was taken as the start of the SWR on that channel. Latency was calculated relative to the time of the SWR of the earliest channel. Channels that did not cross 1(z) were considered maximum latency. The resulting latency image was filtered with a 3 × 3 median filter to remove the effect of bad channels, and upsampled by a factor of 10 for display.

Whole-cell patch-clamp recordings of DVR and claustrum neurons

Long-shank patch pipettes (6–8 MΩ) were pulled from borosilicate glass with a Sutter P1000 electrode puller. Pipettes were filled with internal solution (140 mM K-gluconate, 4 mM NaCl, 14 mM phosphocreatine, 10 mM HEPES, 4 mM Mg-ATP, 0.3 mM Na-GTP, 4 mg ml⁻¹ biocytin). Experiments were carried out on an upright Olympus BX61WI microscope with 5× and 40× water-immersion objectives and cells were patched under visual guidance. Excitatory and inhibitory postsynaptic currents were recorded in the voltage-clamp configuration with the same cell held at either -70 mV or +10 mV. Simultaneous patch-clamp and LFP recordings were carried out with an EPC10 Quadro amplifier (HEKA).

Pharmacology

Serotonin hydrochloride (0.1–30 µM), carbamoylcholine chloride (50 µM), noradrenaline bitartrate (25 µM), SKF38393 hydrobromide (10 µM), (R)-(+)-8-hydroxy-DPAT hydrobromide (2 µM), L-703,664 succinate (1 µM), CP 809101 hydrochloride (0.1 µM), LP44 (0.2 µM) and TTX (20 µM) were diluted to their final concentrations in ACSF (126 mM NaCl, 3 mM KCl, 1.8 mM CaCl_2 , 1.2 mM MgCl_2 , 24 mM NaHCO_3 , 0.72 mM

NaH_2PO_4 , 20 mM glucose, pH 7.4). For slice experiments, drugs were continuously bath-applied after a baseline recording period of 5–20 min. For ex vivo experiments in Extended Data Fig. 8, TTX dissolved in ACSF was injected into the claustrum through a glass micropipette using a 10-ml syringe pressurizer (20–30 hPa for 15 min). For serotonin uncaging, RuBi-5HT (Abcam) (10 µM) was bath-applied, and white light (400–700 nm, 0.11 W cm⁻², TH4-200, Olympus) was turned on and off at chosen intervals (for example, 80 s).

We tested several metabotropic 5HT₁ agonists. Of those, the 5HT_{1D} agonist L-703,664 best mimicked the effects of serotonin, consistent with the high expression of 5HT_{1D} in glutamatergic neurons in the claustrum (Extended Data Fig. 5a). The 5HT₁₇ agonist LP44 had no effect (Fig. 4f), which is also consistent with the low expression of 5HT₁₇ in claustrum excitatory neurons. The 5HT_{2C} agonist CP 809101 increased the rate but not the amplitude of SWRs.

scRNA-seq libraries

Adult male lizards (150–400 g) were deeply anaesthetized with isoflurane, ketamine (50 mg kg⁻¹) and midazolam (0.5 mg kg⁻¹) and decapitated. The head was immersed in ice-cold oxygenated ACSF (126 mM NaCl, 3 mM KCl, 2 mM CaCl_2 , 4 mM MgCl_2 , 24 mM NaHCO_3 , 0.72 mM NaH_2PO_4 , 20 mM glucose, pH 7.4). The brains were perfused to remove blood from the vasculature. The data shown originate from four libraries constructed from data from one male lizard (160 g, 20 months old).

Thereafter, the brain was removed and immersed in oxygenated ice-cold ACSF. The brain was embedded in 4% low-melting agarose, glued to the base of a vibratome (VT1200S, Leica) and immersed in ice-cold oxygenated ACSF, and 500-µm-thick sections were prepared (speed, 0.08 mm s⁻¹). The sections were individually inspected under a dissection microscope (Stemi 2000-C, Zeiss) and anatomical regions of interest were dissected (telencephalon, amDVR). These slices were cut with fine scissors (Fine Science Tools) into small cubes of tissue (around 500 × 500 × 500 µm).

These were transferred to dissociation buffer (20 U ml⁻¹ papain, 200 U ml⁻¹ DNase I, 25 µg ml⁻¹ liberase TM, 1 µM TTX, 100 µM D-APV) and triturated with fire-polished, silanized glass pipettes of decreasing tip diameter (around 10 passes per pipette). After every pipette change the supernatant (dissociated cell suspension) was removed and filtered through a strainer with 100-µm mesh.

The pooled dissociated cell suspension was diluted to 20 ml (with Hibernate A – CaCl_2), transferred to a 50-ml reaction tube and filtered with a strainer with 40-µm mesh. Then 5 ml of 4% bovine serum albumin (BSA) in Hibernate A – CaCl_2 was added to the bottom of the tube with a long-stemmed glass pipette. The solution was spun in a centrifuge at 300g at 4 °C (lowest acceleration and brake) for 5 min. The supernatant was removed and the cell pellet resuspended in 20 ml of Hibernate A – CaCl_2 . This procedure was repeated for a second-gradient clean-up. The pellet was then resuspended in an appropriate amount (50–200 µl) of Hibernate A – CaCl_2 – MgCl_2 and the cell concentration was measured with a Fuchs–Rosenthal cell-counting chamber (Brand).

The cell suspension was then diluted to 466 cells µl⁻¹ and used as input to half a chip (four samples) of the 10X Chromium system (Chemistry v.3) with a targeted cell recovery of 7,000 cells per sample. The library construction was performed according to the manufacturer's instructions.

The final four libraries were quantified using Qubit fluorometer (Thermo Fisher Scientific) and sequenced five times on a DNA sequencer (NextSeq 500, Illumina) with an average depth of 442,806,563 reads per library.

Analysis of transcriptomics data

Raw sequencing data were processed using Cellranger v.3.0 (10X Genomics). Raw reads were demultiplexed and filtered with the cellranger mkfastq function with default settings. To generate digital gene expression matrices, demultiplexed reads were aligned to the *Pogona*

genome with the cellranger count function, setting the force-cells parameter to 7,000. For reads alignment, we reannotated the *Pogona* genome (assembly 1.1.0, NCBI accession number GCF_900067755.1, 10 April 2017) using the same 3'-end MACE (massive analysis of cDNA ends) data and the approach described previously¹¹.

Digital gene expression matrices were analysed in R, using the Seurat v.3.0 package¹⁴. Cells were filtered by number of genes (more than 800 genes per cell) and percentage of mitochondrial genes (lower than 5%), yielding a total of 20,257 cells, with a median number of 2,278 transcripts and 1,349 genes per cell. Data were normalized by the total number of transcripts detected in each cell, and regressed by the number of genes and of transcripts (by setting `vars.to.regress = c("nFeature_RNA", "nCount_RNA")` in ScaleData function). Variable genes were identified after variance standardization from an estimate of the mean variance relationship (FindVariableFeature, method = "vst"), and the top 1,000 highly variable genes were used for principal component analysis. The first 30 principal components were used for Louvain clustering (FindClusters, resolution = 0.2) and for dimensionality reduction with UMAP⁴² (RunUMAP with default settings).

After this first round of analysis, neuronal clusters (characterized by high expression of pan-neuronal markers, such as the synaptic protein *snap25*) were analysed again using the above procedure with the following settings: more than 800 genes per cell, 2,000 highly variable genes, 30 principal components, clustering resolution = 0.2. This led to the identification of 12 neuronal clusters. One cluster of doublets, recognized by the co-expression of markers of glutamatergic and GABAergic (γ -aminobutyric acid-producing) neurons, was removed, leaving 9,777 neurons. These were analysed again with the same parameters (but clustering resolution = 2) to yield 33 clusters (Extended Data Fig. 3).

From this neuronal dataset, we identified 4,054 pallial glutamatergic neurons (with more than 1,000 genes per cell) that co-expressed the vesicular glutamate transporters *slc17a7* and *slc17a6*. Further subclustering of these cells (analysis settings: 2,000 highly variable genes, 34 principal components, clustering resolution = 3) led to the identification of 29 clusters (Fig. 3a, Extended Data Fig. 3). To assign an identity to each of these clusters, we analysed the expression of marker genes with known tissue expression patterns¹¹. This allowed us to define the pallial region to which each cluster belongs (for example, hippocampus for *zbtb20*-expressing clusters). Further annotation of cluster identities (Extended Data Fig. 3) was based on the expression of selective markers or combinations of marker genes, identified from the transcriptomics data. Note about gene nomenclature, all reptilian gene names are in lower case in the Article as per *Nature* style; however, in the extended data figures, reptilian gene names are in uppercase, according to convention.

Analysis of ion-channel and neurotransmitter-receptor genes

We mined the *Pogona* genome for the following gene families: noradrenaline, acetylcholine, serotonin and dopamine receptors; calcium, chloride, sodium and potassium channels; and GABA, glutamate, adenosine, cannabinoid, glycine and histamine receptors. This yielded 270 genes in total. Of these, 143 were kept for further analysis, because they were detected in at least 20% of the cells of at least one glutamatergic cluster (Extended Data Fig. 5a).

To calculate pairwise cluster correlations (Pearson correlations, Extended Data Fig. 5b), we used this set of 143 genes and average cluster expression data (calculated from normalized and log-transformed data with the AverageExpression function in the Seurat package). A distance matrix was calculated from the correlation matrix, and used for hierarchical clustering (R package hclust) with the Ward.D2 linkage method.

The gene expression matrix from above was transposed to calculate gene–gene correlations. The gene dendrogram was also calculated with hierarchical clustering and the Ward.D2 linkage method.

The heat map in Fig. 3h was generated from the matrix of 29 glutamatergic clusters (columns) and average expression of the 143 genes (rows). The data matrix was scaled by columns, and the heat map was plotted with the heatmap.2 function from the R package gplots. The dendrogram of glutamatergic clusters is based on Euclidean distance and Ward.D2 linkage.

Mapping of single-cell transcriptomes across species

To map *Pogona* single-cell transcriptomes on mouse single-cell data, we used the dataset from a previous study¹³, which is available on the dropviz.org website. In this dataset, pallial glutamatergic neurons were sampled from three regions: 'hippocampus', 'frontal cortex' and 'posterior cortex'. These dissections encompass several cell types. For example, 'frontal cortex' includes the claustrum, and 'hippocampus' includes the subiculum and entorhinal cortex. Raw data were processed through the Seurat pipeline (normalization, scaling, selection of variable genes) and glutamatergic clusters and subclusters were selected, according to the cluster and subcluster identities provided previously (ref. ¹³ and dropviz.org). Subclusters were downsampled to a maximum number of 200 cells per subcluster, yielding a total of 17,455 cells.

Comparative analysis of *Pogona* and mouse was limited to one-to-one orthologues, according to the orthology annotations provided by Ensembl (*Pogona* assembly pv1.1 and mouse assembly GRCm38.p6, one-to-one orthologues downloaded on 1 May 2019). Of 13,273 one-to-one orthologues, 10,693 were detected in both the mouse and *Pogona* datasets and used for the comparative analysis.

The *Pogona* and mouse data were analysed jointly following the approach described previously¹⁴. In brief, after normalization and scaling, 1,500 highly variable genes were identified in each dataset. The union of these sets of variable genes was used for a joint canonical correlation analysis (CCA). The first 15 canonical components were then used to identify 2,626 transfer anchors; that is, pairs of cells with matching neighbourhoods (mutual nearest neighbours) in the two transcriptomics spaces (function FindTransferAnchors from Seurat). These anchors were then used to project *Pogona* cells (query dataset) on the mouse dataset (reference dataset), using the TransferData function from Seurat. The projection is based on a weighted classifier that assigns a classification score on the basis of the distance of each cell from the transfer anchors. Figure 3g represents the result of the classification, showing the fraction of single cells from each *Pogona* cluster that map to each of the mouse subclusters (mouse subclusters without matching lizard cells are not indicated in the figure).

The approach described above was also used to project the transcriptomes of turtle pallial glutamatergic cells onto the *Pogona* data (Extended Data Fig. 7a). The turtle data are from a previous study¹¹. The comparison was based on 9,820 one-to-one orthologues detected in both species. For this analysis, the top 2,000 variable genes of each dataset were used for CCA. The first 25 canonical components were used to compute 3,406 transfer anchors.

Identification of *Pogona* brain areas with a potential role in brain-state regulation

Areas known to play a part in controlling brain state have been, over the past decades, identified in a number of mammalian species. Those areas can be identified by their location (for example, within the hypothalamus, midbrain or brainstem), their axonal projections and the neuroactive substances that their neurons contain and release (and thus potential marker genes). To our knowledge, no such description exists at present for the brain of the bearded dragon (*Pogona*), but anatomical studies of homologous areas have been performed in other reptilian species^{43–55}. These references were used to identify relevant brain areas, including preoptic area, supramammillary nucleus⁵⁶ and tuberomammillary nucleus in the hypothalamus; ventral tegmental area, substantia nigra and periaqueductal grey in the midbrain; and lateral dorsal tegmental nucleus, locus coeruleus, subcoeruleus and

raphe nucleus in the brainstem. The location and identity of these areas were established in *Pogona* by immunohistochemistry and/or FISH using appropriate neuronal markers, combined with Nissl stains of brain sections. Tyrosine hydroxylase (a marker of catecholaminergic neurons) was used to identify preoptic area, ventral tegmental area, substantia nigra, periaqueductal grey and locus coeruleus; choline acetyltransferase was used to identify lateral dorsal tegmental nucleus; histamine was used to identify tuberomammillary nucleus; serotonin was used to identify raphe nucleus; and subcoeruleus identification was based on the prior identification of lateral dorsal tegmental nucleus and locus coeruleus and by the expression of *slc17a6* (vesicular glutamate transporter 2; a marker of glutamatergic neurons) by in situ hybridization (Extended Data Fig. 6a). The expression of *slc17a6* by in situ hybridization was also used for the identification of supramammillary nucleus⁵⁶ (Extended Data Fig. 6a).

Pogona whole-brain images

Pogona brain reconstruction (Fig. 3i) was based on images obtained with a μ CT scanner, and the 'surface' function of the Imaris software (Oxford Instruments). The boundaries of relevant nuclei were determined from consecutive serial histological sections. The serial images were aligned and assembled to three-dimensional (3D) volumes using the Voloom software, and then imported into Imaris and aligned with the 3D data. The boundaries of some areas identified by retrograde tracing were defined from GFP and Nissl staining patterns.

Immunohistochemistry and in situ hybridization

The lizards were deeply anaesthetized with isoflurane, ketamine (60 mg kg⁻¹) and midazolam (2 mg kg⁻¹) until loss of the foot-withdrawal reflex. Pentobarbital (10 mg kg⁻¹) was then administered by intraperitoneal injection. After loss of the corneal reflex, the lizard was perfused transcardially with cold PBS (1.47 \times 10⁻³ M KH₂PO₄, 8.10 \times 10⁻³ M Na₂HPO₄·12H₂O, 2.68 \times 10⁻³ M KCl, 1.37 \times 10⁻¹ M NaCl) followed by 4% paraformaldehyde (PFA) in PBS. The brain samples were post-fixed with 4% PFA–PBS for 16 h at 4 °C and subsequently immersed in 30% sucrose for 24 h at 4 °C. The brain area was sectioned coronally (60 μ m) with a microtome at –24 °C. The sections were permeabilized for 30 min at room temperature in blocking solution (PBST: PBS with 0.3% Triton X-100 and 10% goat serum) and incubated with primary antibodies (anti-GFP, A10262, Invitrogen, chicken, 1:1,000; hippocalcin, ab24560, Abcam, rabbit, 1:1,000; ChAT, AB144P, Merck, goat, 1:100; mTH, 22941, ImmunoStar, mouse, 1:100; rabTH, AB152, Merck, rabbit, 1:200; histamine, 22939, ImmunoStar, rabbit, 1:100; serotonin, MAB352, Merck, rat, 1:100) in blocking solution overnight at 4 °C. After washing with PBST three times, the samples were incubated with secondary antibodies conjugated with appropriate secondary antibodies (1:500, all from Invitrogen) in blocking solution for 4 h at room temperature, followed by three washes with PBST. Some slices were counterstained with NeuroTrace 435/455 blue-fluorescent Nissl stain (N21479, Invitrogen, 1:200) in PBS for 2 h at room temperature. After rinsing with PBS, the samples were mounted with Dako Fluorescence Mounting Medium (S3023, Dako) or Roti-Mount FluorCare DAPI (HP20.1, Carl Roth). Images were acquired using a confocal system or fluorescent microscopy at 10 \times , 20 \times or 40 \times . Chromogenic in situ hybridization and dual colorimetric in situ hybridization were performed following the protocols previously described¹¹.

Fluorescent in situ hybridization by RNAscope

The lizards were deeply anaesthetized as described above. After loss of corneal reflex, the animals were killed by decapitation. Brains were dissected out immediately, embedded in OCT on a dry-ice ethanol bath and stored at –80 °C. Fresh frozen brains were sectioned at 25 μ m on a Thermo Fisher Scientific CryoStar NX70 cryostat and placed onto SuperFrost-coated (Thermo Fisher Scientific) slides. Some slides were stored at –80 °C after air-drying. RNAscope hybridization was

performed according to the manufacturer's instructions. We used the RNAscope Multiplex Fluorescent assay (Advanced Cell Diagnostics) for fresh-frozen sections. Target genes and probe catalogue numbers were Pv-CHAT-C2, 522631-C2; Pv-SLC17A6-C1, 529431-C1. Fluorescent Nissl was used for counterstaining. Slides were mounted with ProLong Gold Diamond Antifade Mountant (P36970, Thermo Fisher Scientific). Images were acquired with a digital slide scanner (Pannoramic MIDI II, 3DHISTECH) at 20 \times magnification.

Tract tracing

The lizards were anaesthetized as described for in vivo recordings. Extensive preliminary searches for useful AAV serotypes for reptilian brains and for appropriate incubation conditions were carried out by L. Pammer²². The tracers (rAAV2-retro-CAG-GFP, 37825-AAVrg; rAAV2-retro-hSyn-EGFP, 50465-AAVrg; AAV9-CB7.C1.mCherry.WPRE.RBG, 105544-AAV9; all from Addgene, <https://www.addgene.org>) were injected into one or two forebrain locations (for example, dorso-medial cortex, DVR, amDVR, and so on). Four to six weeks later, the animals were deeply anaesthetized as described above, and after loss of corneal reflex, the animals were killed by decapitation. Brains were dissected out, processed for histology, sectioned and imaged. The data presented come from 18 of 30 injected brains. The remaining 12 brains were rejected either because the viral injections failed or because the injections were not sufficiently specific. Targeting specific regions in the brain of *Pogona* and *Trachemys* is difficult because the brain is loosely contained in the cranial cavity and its position relative to the cranium and reliable landmarks is thus variable: the brain floats in CSF, attached by cranial nerves. As a consequence, there are no reliable stereotactic coordinates based on cranium landmarks. The lateral ventricles are large. The external appearance of the forebrain also lacks reliable landmarks (for example, blood vessels or sulci). Finally, these animals are not standardized species, bred over generations to reduce variability.

Note that, because rAAV2-retro does not infect all neuronal types equally²³, the results from negative retrograde labelling should be confirmed with other methods. Conversely, the connectivity estimated with the tracers we used is likely to be underestimated.

Statistics and reproducibility

Unless stated otherwise, data are mean \pm s.e.m. For comparisons of two groups we performed a two-tailed unpaired *t*-test, two-tailed paired *t*-test, Mann–Whitney rank-sum test or Wilcoxon signed-rank test, as appropriate (all two-sided). For multiple comparisons we performed a Bonferroni test. Significance was determined at the 0.05 α level for all statistical tests. For box plots (Fig. 4d): margins are 25th and 75th percentiles; red, median; whiskers, boundaries before outliers; outliers (+) are values beyond 1.5 \times interquartile range from the box margins. Experiments were repeated independently several times with similar results, with numbers of repetitions as follows: Fig. 1b–e: 7 times; Fig. 2a–c: 4 times; Fig. 2e: 13 times (amDVR) and 9 times (pDVR); Fig. 3: 4 times (a, b, d–h) and 10 times (c); Fig. 4a–c: 3 times; Extended Data Fig. 1a–d: 7 times; Extended Data Fig. 1h: 3 times; Extended Data Fig. 2b: 15 times; Extended Data Fig. 2g, i: 12 times; Extended Data Fig. 2h, j: twice; Extended Data Fig. 3f: 3 times; Extended Data Fig. 4a: 3 times; Extended Data Fig. 4b: 13 times (amDVR) and 9 times (pDVR); Extended Data Fig. 6a–c: 3 times (for all except for c5–7, for which experiments were reproduced once in 5 experiments) (see Fig. 3 legend); Extended Data Fig. 7b: 3 times; Extended Data Fig. 7c: 5 times; Extended Data Fig. 7d: 4 times; Extended Data Fig. 7e, g: 3 times; Extended Data Fig. 8a, e, f: 4 times; Extended Data Fig. 8b–d: 4 times; Extended Data Fig. 9a–d: twice (a, b) and 3 times (c) (claustrum lesions (d) were confirmed in all these experiments); Extended Data Fig. 10a, b: twice (a) and 3–4 times (b).

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Sequencing data have been deposited in the NCBI Sequence Read Archive: BioProjects PRJNA591493 (lizard) and PRJNA408230 (turtle). Links to those archives and to analysis code can be found at: <https://brain.mpg.de/research/laurent-department/software-techniques.html>. Data are also available from the corresponding author on request.

43. Moreno, N., Domínguez, L., Morona, R. & González, A. Subdivisions of the turtle *Pseudemys scripta* hypothalamus based on the expression of regulatory genes and neuronal markers. *J. Comp. Neurol.* **520**, 453–478 (2012).
44. Medina, L., Smeets, W. J., Hoogland, P. V. & Puellas, L. Distribution of choline acetyltransferase immunoreactivity in the brain of the lizard *Gallotia galloti*. *J. Comp. Neurol.* **331**, 261–285 (1993).
45. Bruce, L. L. & Neary, T. J. Afferent projections to the ventromedial hypothalamic nucleus in a lizard, *Gekko gekko*. *Brain Behav. Evol.* **46**, 14–29 (1995).
46. Bruce, L. L. & Neary, T. J. Afferent projections to the lateral and dorsomedial hypothalamus in a lizard, *Gekko gekko*. *Brain Behav. Evol.* **46**, 30–42 (1995).
47. Ebner, F. F. in *Evolution of Brain and Behavior in Vertebrates* (eds Masterton, R. B. et al.) 115–167 (Taylor & Francis, 1976).
48. Font, C., Lanuza, E., Martínez-Marcos, A., Hoogland, P. V. & Martínez-García, F. Septal complex of the telencephalon of lizards: III. Efferent connections and general discussion. *J. Comp. Neurol.* **401**, 525–548 (1998).
49. Hoogland, P. V. & Vermeulen-Vanderzee, E. Efferent connections of the dorsal cortex of the lizard *Gekko gekko* studied with *Phaseolus vulgaris*-leucoagglutinin. *J. Comp. Neurol.* **285**, 289–303 (1989).
50. Smeets, W. J. & Steinbusch, H. W. Distribution of noradrenaline immunoreactivity in the forebrain and midbrain of the lizard *Gekko gekko*. *J. Comp. Neurol.* **285**, 453–466 (1989).
51. Smeets, W. J., Hoogland, P. V. & Voorn, P. The distribution of dopamine immunoreactivity in the forebrain and midbrain of the lizard *Gekko gekko*: an immunohistochemical study with antibodies against dopamine. *J. Comp. Neurol.* **253**, 46–60 (1986).
52. ten Donkelaar, H. J., Bangma, G. C., Barbas-Henry, H. A., de Boer-van Huizen, R. & Wolters, J. G. The brain stem in a lizard, *Varanus exanthematicus*. *Adv. Anat. Embryol. Cell Biol.* **107**, 1–2 (1987).
53. ten Donkelaar, H. J. in *The Central Nervous System of Vertebrates* Vol. 1–3 (eds Nieuwenhuys, H. et al.) 1315–1524 (Springer, 1998).
54. Wolters, J. G., ten Donkelaar, H. J., Steinbusch, H. W. & Verhofstad, A. A. Distribution of serotonin in the brain stem and spinal cord of the lizard *Varanus exanthematicus*: an immunohistochemical study. *Neuroscience* **14**, 169–193 (1985).
55. Wolters, J. G., ten Donkelaar, H. J. & Verhofstad, A. A. Distribution of catecholamines in the brain stem and spinal cord of the lizard *Varanus exanthematicus*: an immunohistochemical study based on the use of antibodies to tyrosine hydroxylase. *Neuroscience* **13**, 469–493 (1984).
56. Pedersen, N. P. et al. Supramammillary glutamate neurons are a key node of the arousal system. *Nat. Commun.* **8**, 1405 (2017).

Acknowledgements We thank K. Steele and A. Schwartzlose for help with serotonin immunocytochemistry and neural tract tracing; E. Northrup and G. Wexel for veterinary care; T. Klappich for reptile care; S. Junek, F. Vollrath and C. Polisseni for help with imaging and microscopy; G. Tushev for help with genome reannotation; E. Desfilis for help with lizard neuroanatomy; and members of the Laurent laboratory for help and comments. The work was funded by the Max Planck Society, the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) (ERC grant agreement no. 322705), the European Research Council under the European Union's Horizon 2020 research and innovation programme (ERC grant agreement no. 834446) and the DFG (CRC1080) (G.L.); postdoctoral fellowships from the JSPS (SPD and for research abroad) and from the Kanae Foundation for the promotion of medical science (H.N.); and an EMBO long-term fellowship (ALTF 421-2017) (L.A.F.).

Author contributions H.N. and L.A.F. contributed equally and have equal right to list themselves first in bibliographic documents. H.-H.L. and M.A.T. also contributed equally. Project conception, H.N., L.A.F. and G.L.; animal surgery, M.K., H.N. and L.A.F.; electrophysiology, H.N., L.A.F. and S.R.; pharmacology, H.N., R.K. and L.A.F.; scRNA-seq, T.G.F., D.H., A.M. and M.A.T.; bioinformatics, M.A.T. and D.H.; tracing, anatomy and histology, L.A.F., H.-H.L., R.K., T.G.F., A.A. and M.K.; experimental design, data interpretation and analysis, H.N., L.A.F., H.-H.L., M.A.T., T.G.F., D.H., S.R. and G.L.; project management and supervision, G.L.; manuscript writing, G.L., with input from all co-authors.

Competing interests The authors declare no competing interests.

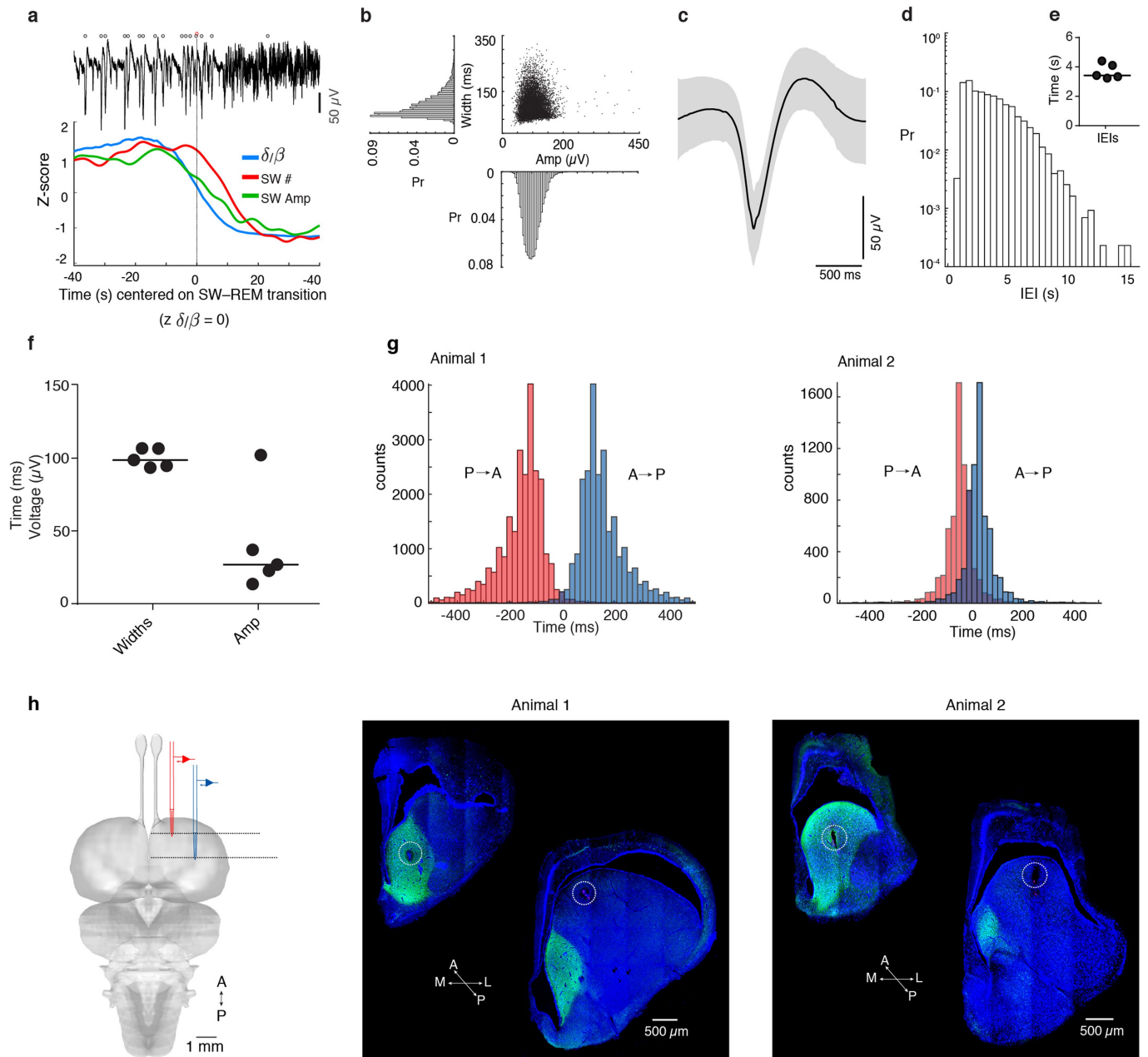
Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-1993-6>.

Correspondence and requests for materials should be addressed to G.L.

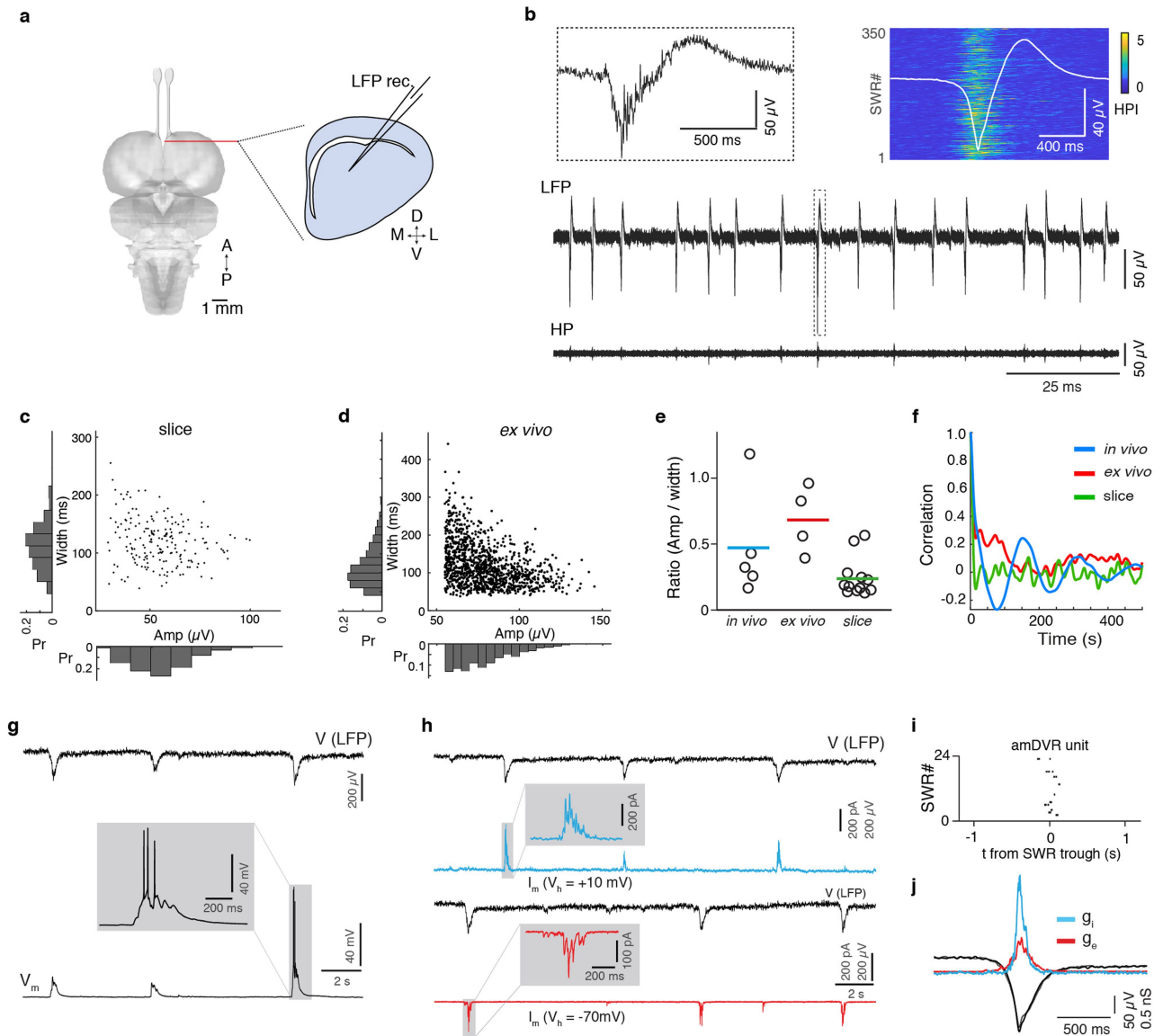
Peer review information *Nature* thanks Michael Lazarus and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



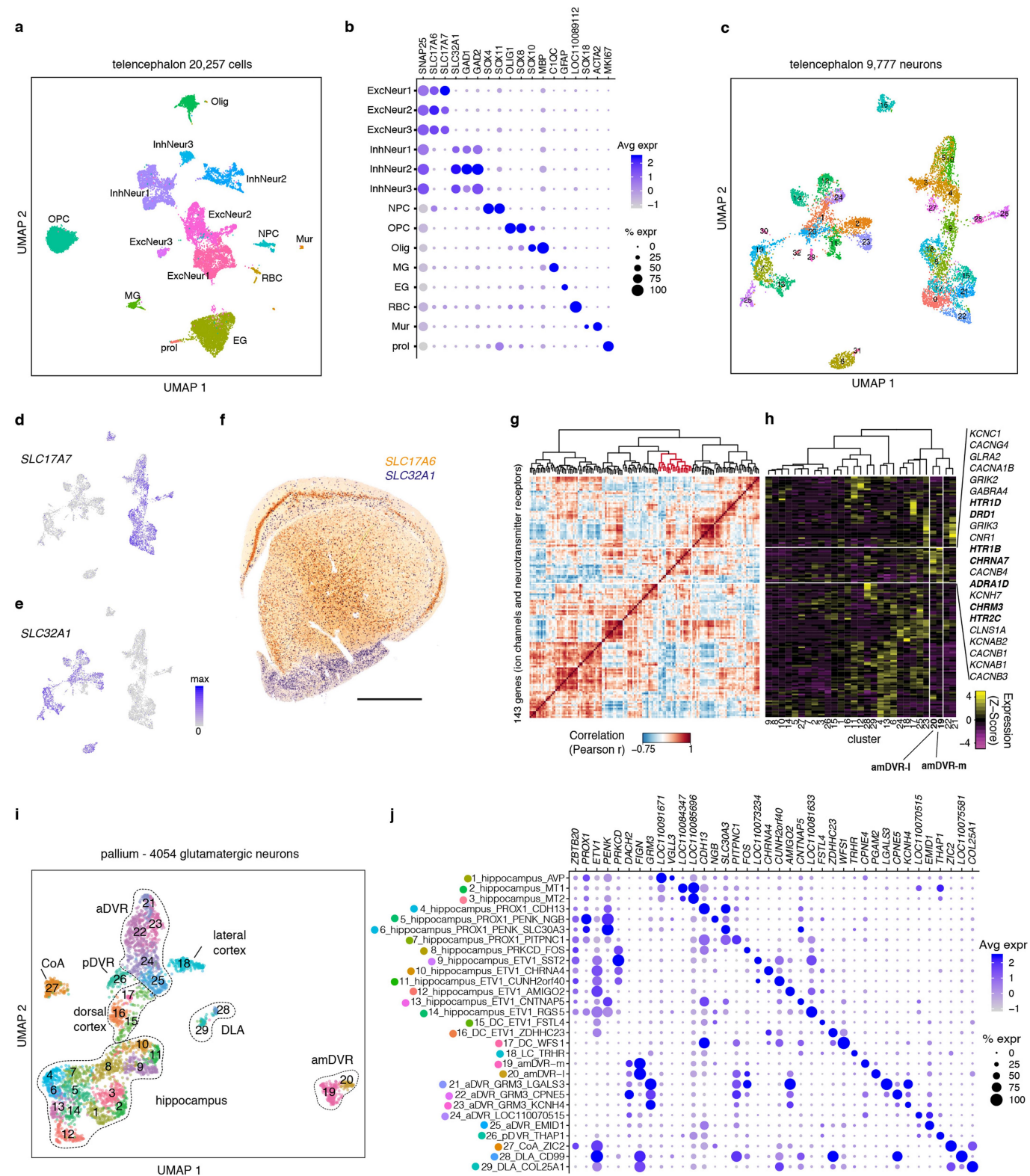
Extended Data Fig. 1 | Further description of SWR statistics and propagation in vivo. **a**, The amplitude and frequency of sharp waves vary as the animal transitions between slow-wave and REM sleep. Top, illustrative LFP trace (<150 Hz) showing a decrease in sharp-wave amplitude and frequency around the slow-wave-REM transition point. Open circles indicate detected sharp waves³ (see Methods). Data in **a-d** are from the same animal and a single night, and correspond to the recording in Fig. 1 (anterior recording site, red). Statistics are based on $n=11,123$ sharp waves. **b**, Distribution of sharp-wave width (measured at half peak amplitude) and peak amplitude from the animal in **a** and Fig. 1. Pr, probability. **c**, Average sharp-wave trace ± 1 s.d. (grey) calculated over $n=11,123$ sharp waves. **d**, Inter-event interval (IEI) for sharp waves recorded during slow-wave sleep. The y axis (probability) is on a logarithmic scale. **e, f**, Summary of data recorded over five nights from two animals. Each

circle represents the mean of one night; black line shows the median. **e**, Mean inter-event intervals during slow-wave sleep. **f**, Mean sharp-wave width and amplitude ($n=8,055-13,494$ sharp waves per night). **g**, Delay distributions of sharp waves in anterior (or posterior) DVR, triggered on simultaneously recorded posterior (or anterior) DVR. Sharp waves from three nights (animal 1; $n=24,501$ sharp waves) and two nights (animal 2; $n=13,070$ sharp waves). **h**, Locations of simultaneous recording sites in the aDVR (circles). Left, schematic of recording configuration. Middle and right, confocal images highlighting the recording sites, as identified by electrolytic lesions and Dil dye that was applied to the back of the silicon probes. Post hoc staining with an antibody against hippocalcin was used to determine the borders of the claustrum (see Fig. 3).



Extended Data Fig. 2 | Comparison of SWR statistics across preparations and recording conditions. **a**, Slice preparation (see Methods) for field-potential recordings. **b**, Spontaneous sharp waves (LFP; <150 Hz) and corresponding ripples (high-pass (HP) band; 70–150 Hz) in the amDVR. Insets: top left, magnification of the SWR marked with a dotted box; top right, 350 ripples; high-pass signal intensity (HPI) >70 Hz aligned on trough of sharp wave (overlaid as average). **c**, Distribution of amplitude (x) and width (y, full width at half maximum) of SWR events in a representative DVR slice. **d**, Distribution of SWR amplitude and width (as in **c**) in a representative ex vivo preparation. **e**, Ratio of amplitude (μ V) to width (ms); $n = 5$ sleep epochs from 3 animals (in vivo; blue), 4 ex vivo brains (red) and 12 slices (green). Lines show the mean. **f**, Autocorrelation function of sharp-wave times, showing that the characteristic rhythmic modulation of sharp-wave generation (which is due to the alternation of slow-wave sleep and REM sleep with a 2–3 min period) in sleeping animals is absent from both ex vivo brain preparations and slice preparations ($n = 5$ sleep epochs from 3 animals (in vivo), 4 ex vivo brains and 12

slices). **g**, Whole-cell patch-clamp recording (in current-clamp mode) of a DVR neuron (V_m), together with LFP recording in a neighbouring region (V (LFP)) with a glass micropipette. Note the simultaneous depolarization of the neuron and SWRs, and moderate neuronal depolarization that gives rise to occasional firing (three action potentials here). The experiment was repeated with 12 neurons. **h**, Whole-cell patch-clamp recording of an amDVR neuron in voltage-clamp mode, held at depolarized (blue) and hyperpolarized (red) holding potentials (V_h). Note the volleys of excitatory (red) and inhibitory (blue) currents at each SWR (LFP), and the near absence of synaptic input in between. **i**, Spike times of a patched amDVR neuron in relation to sharp waves. Note the locking to the sharp-wave trough ($t = 0$), and the absence of firing otherwise ($n = 2$ amDVR neurons). **j**, Mean excitatory (g_e) and inhibitory (g_i) conductances ($n = 20$ and 21 events, respectively). The black and grey lines show averaged sharp waves recorded with inhibitory and excitatory conductances, respectively. Traces are aligned on the sharp-wave trough.

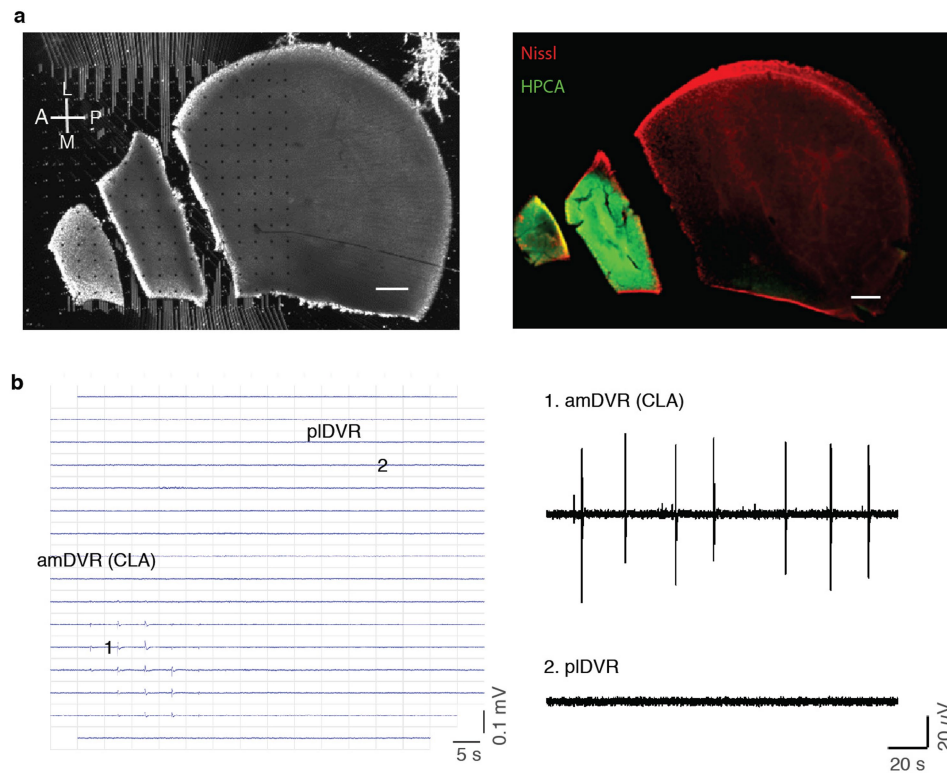


Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Additional single-cell transcriptomic

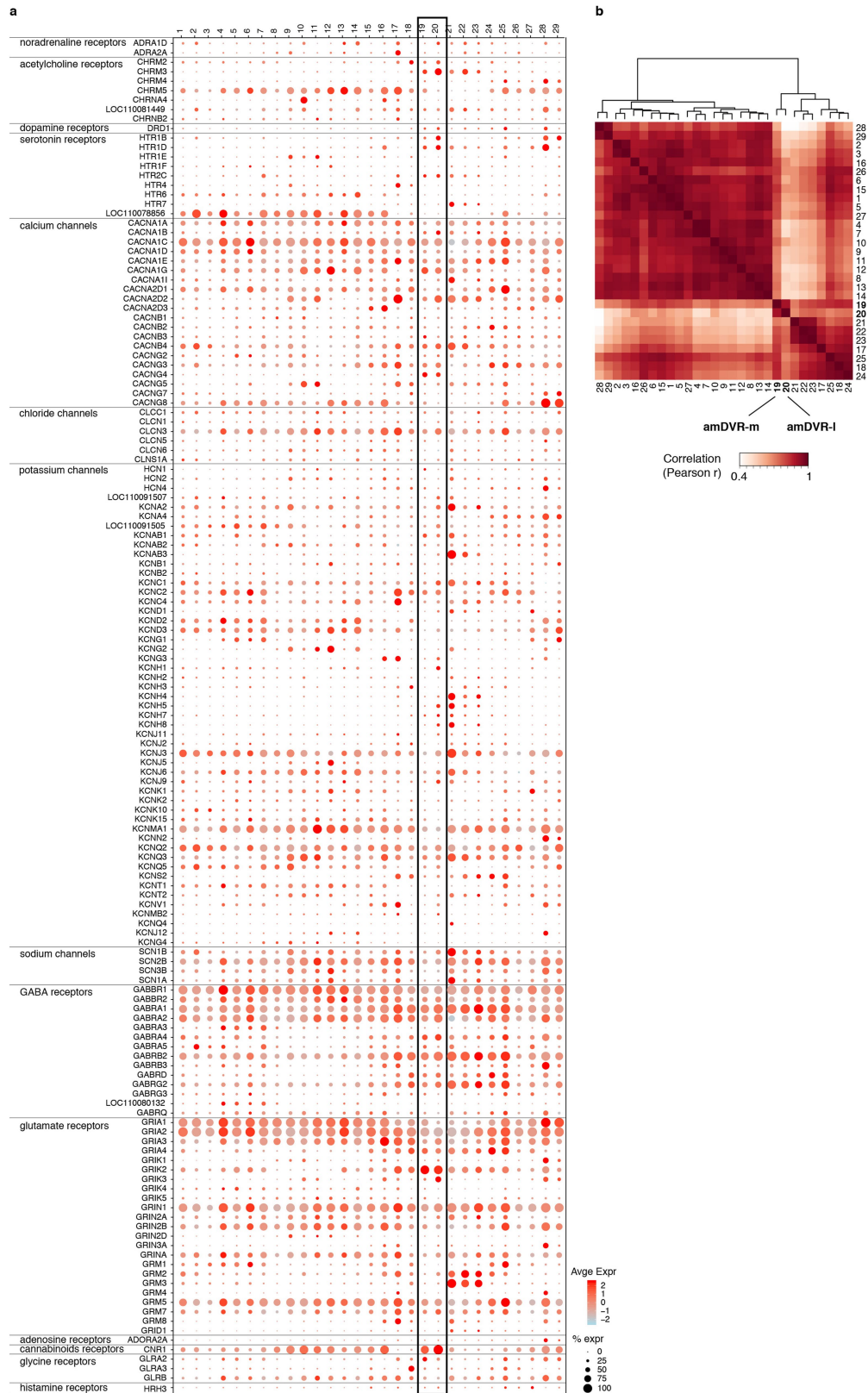
characterization. a, UMAP⁴² representation of 20,257 *Pogona* telencephalic cells, colour-coded by cluster. EG, ependymoglia cells; ExcNeur, excitatory neurons; InhNeur, inhibitory neurons; MG, microglia; mur, mural cells; NPC, neural progenitor cells; olig, oligodendrocytes; OPC, oligodendrocyte progenitor cells; prol, proliferating cells; RBC, red blood cells. **b**, Dot plot showing the expression of canonical cell markers (rows) across telencephalic cell clusters (columns). The size of the dot corresponds to the percentage of cells in a cluster in which the gene has been detected, and the colour represents the expression level. **c**, UMAP representation of 9,777 lizard telencephalic neurons, colour-coded by cluster. **d, e**, UMAP representations of glutamatergic (*slc17a7*) and GABAergic (*slc32a1*) neurons in the telencephalon dataset. **f**, Double colorimetric in situ hybridization in a frontal section through the anterior *Pogona* forebrain. Scale bar, 1 mm. *slc32a1* (blue) labels GABAergic neurons in the subpallium and scattered GABAergic neurons that have

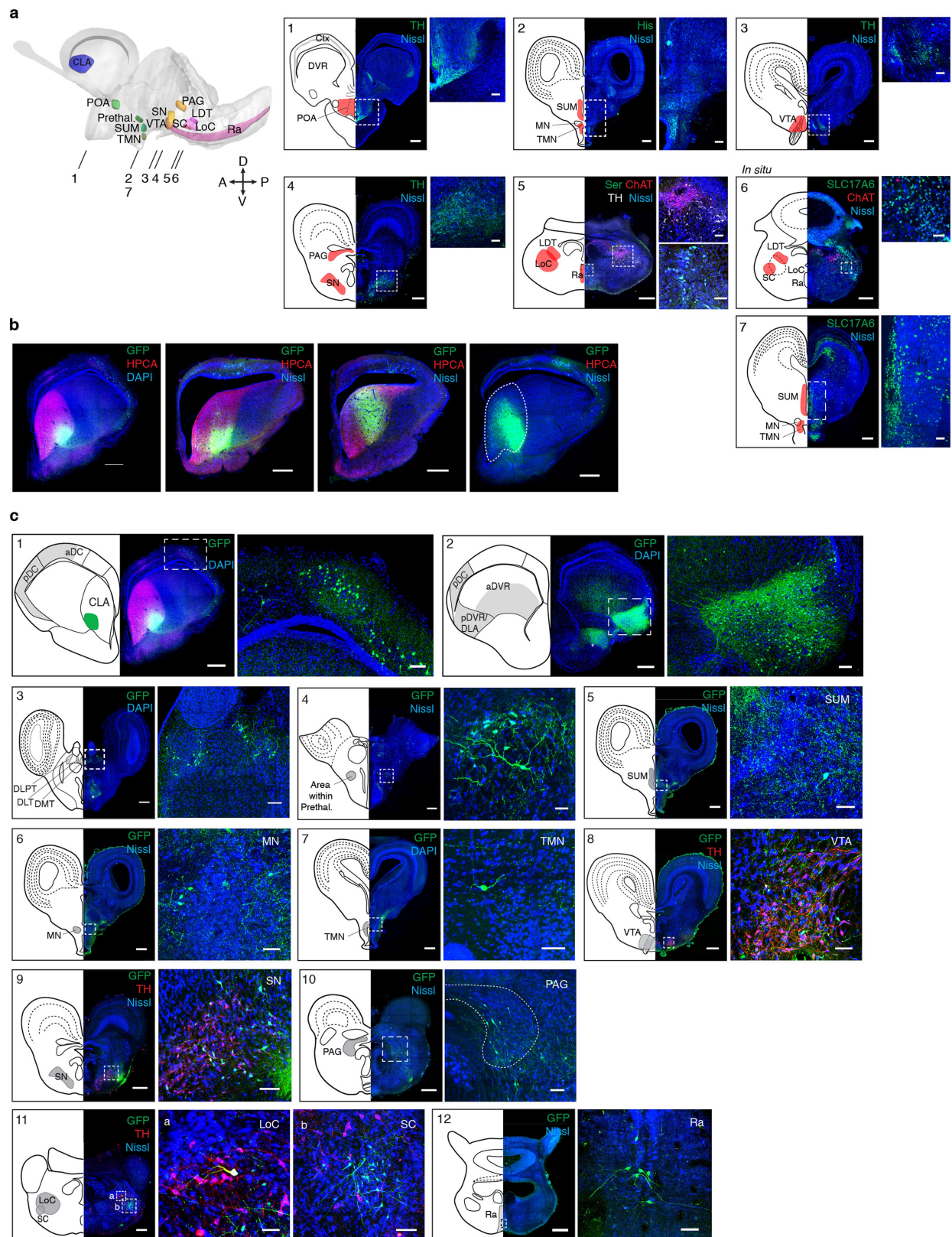
migrated from subpallium to pallium. *slc17a6* (orange) labels glutamatergic neurons in the pallial region. **g**, Ordered matrix of pairwise Pearson correlations between the expression of 143 ion-channel and neurotransmitter-receptor genes detected in this glutamatergic pallial dataset from *Pogona* (see Extended Data Fig. 5). The dendrogram (top) is based on correlation coefficients and Ward.D2 linkage; red indicates a gene module with enriched expression in the amDVR. **h**, Average expression, in the 29 glutamatergic *Pogona* clusters, of the 143 genes in **g** (and Extended Data Fig. 5). Genes with enriched expression in the amDVR are listed on the right, with relevant neurotransmitter receptor genes in bold. **i**, UMAP representation of 4,054 lizard pallial glutamatergic neurons, colour-coded by cluster (same as in Fig. 3a). **j**, Dot plot showing the expression of specific cluster markers (rows) in the 29 pallial glutamatergic clusters (columns). The size of the dot corresponds to the percentage of cells in a cluster in which the gene has been detected, and the colour represents the expression level.



Extended Data Fig. 4 | Mini-slices of the DVR and localization of SWR generation. **a**, Left, recording configuration of mini-slices of the DVR on a planar 252-channel microelectrode array. Dots represent electrodes. Right, post hoc immunostaining of the mini-slices. Red, Nissl; green, hippocalin. **b**, Left, spatial distribution of SWR waveforms as recorded from the mini-slices

in **a**. Right, illustrative LFP traces recorded from the amDVR or claustrum (1) and pIDVR (2) (see recording positions on the microelectrode array on the left). In conclusion, SWRs occur spontaneously in the amDVR, and are absent from the pIDVR once it is disconnected from the amDVR (claustrum).

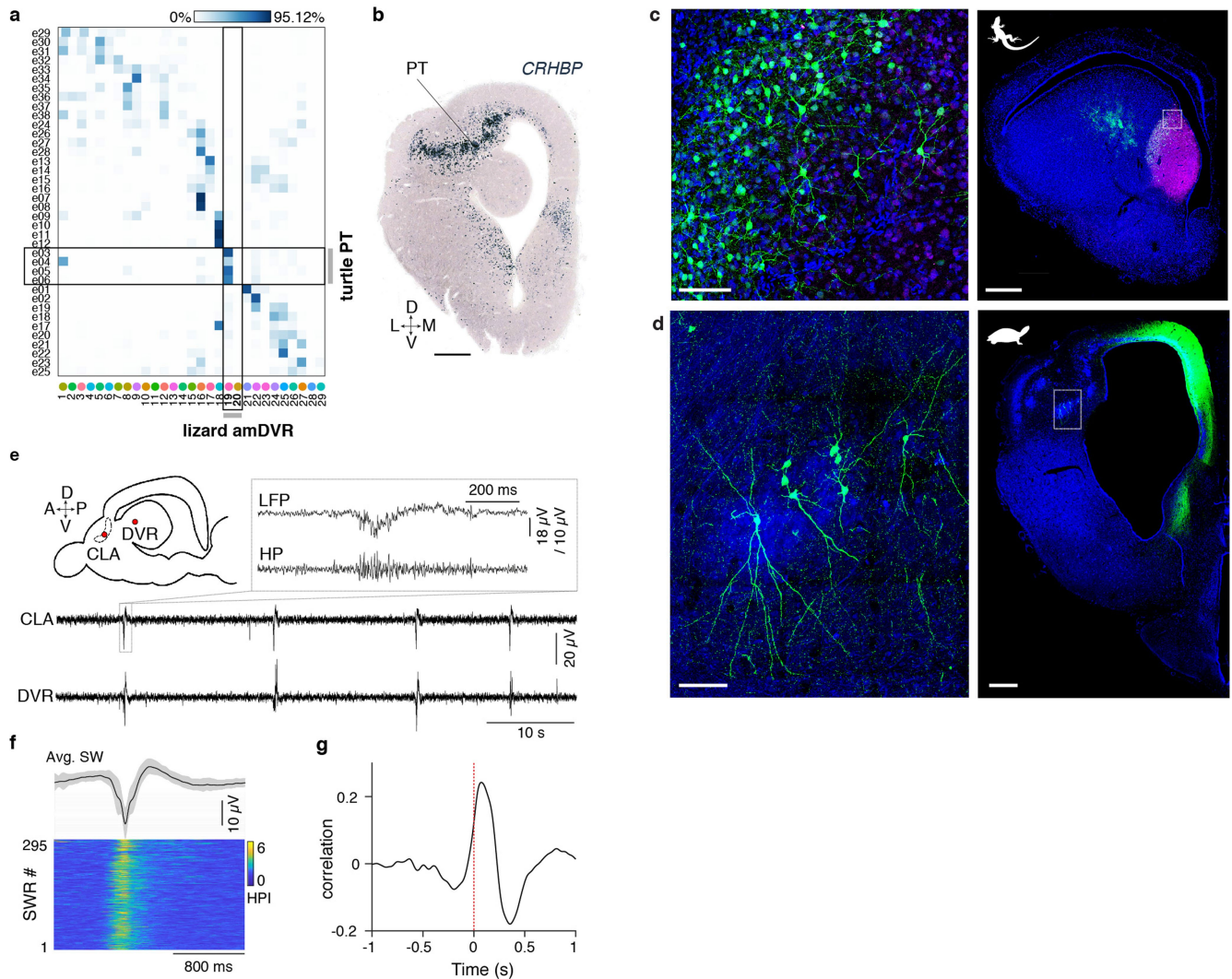




Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Identification of potential regulatory areas of brain states and distribution of GFP-labelled neurons after injection of rAAV2-retro into the claustrum. **a**, Left, schematic of the *Pogona* brain in sagittal view, showing the regions defined by immunohistochemistry, in situ hybridization and retrograde tracing. Numbers 1–7 indicate the levels of the transverse sections that are shown on the right. Right (panels 1–7), micrographs and corresponding schematic representations of relevant areas (in red), identified by immunohistochemistry, in situ hybridization and Nissl staining. Scale bars, 500 μ m. Far right of panels 1–7, magnified views of area(s) delineated as box(es) in the corresponding photomicrographs. Scale bars, 100 μ m. **b**, Identification of rAAV2-retro injection sites. Scale bars, 500 μ m. The red channel is not shown in the rightmost image. **c**, Illustrative examples of retrograde labelling of claustrum connectivity, in transverse sections. Panels 1,

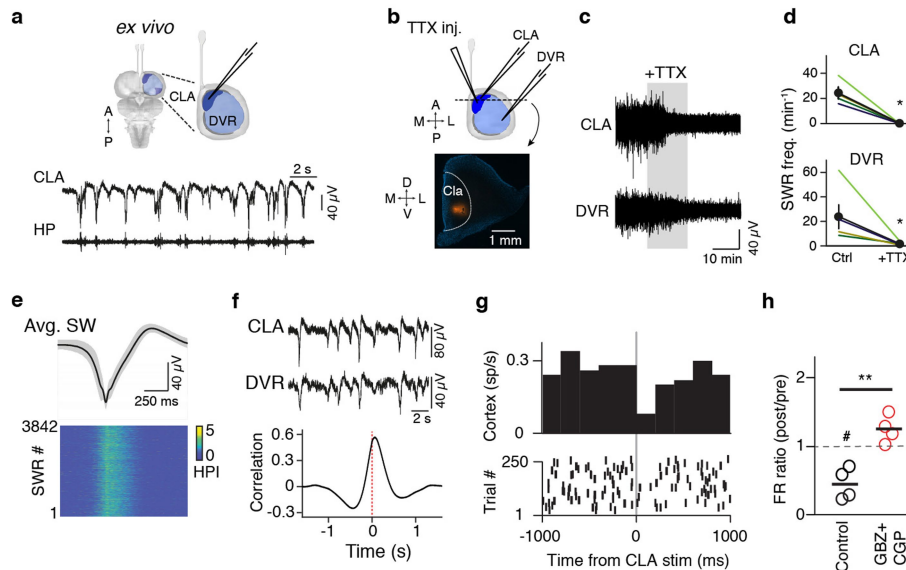
2, inputs to claustrum revealed by rAAV2-retro injection in the claustrum. Panel 1, injection site in lateral claustrum. The claustrum is indicated by anti-hippocalcin immunostain (pink). Note retro-labelled cells in the anterior dorsal cortex (box, magnified at right). Panel 2, same brain as in 1, but a more posterior section. The labelled region in the box is the dorsal lateral amygdala. Panels 3–12, representative images illustrating the distribution of GFP-labelled neurons in the DLPT, DLT, DMT, prethalamus, SUM, mammillary nucleus (MN), TMN, VTA, SN, PAG, LoC and SC, with projections to the claustrum. Abbreviations as in Fig. 3. The catecholaminergic neuron marker tyrosine hydroxylase (TH) was used to indicate the location of the VTA, SN and LoC. Scale bars, 500 μ m. Scale bars for magnified areas: DLPT, DLT, DMT, prethalamus, SUM, MN, TMN, VTA, LoC, 50 μ m; SN, PAG, SC, 100 μ m.



Extended Data Fig. 7 | The claustrum of lizard and turtle differ in position and architectonics, but are both autonomous sources of SWRs.

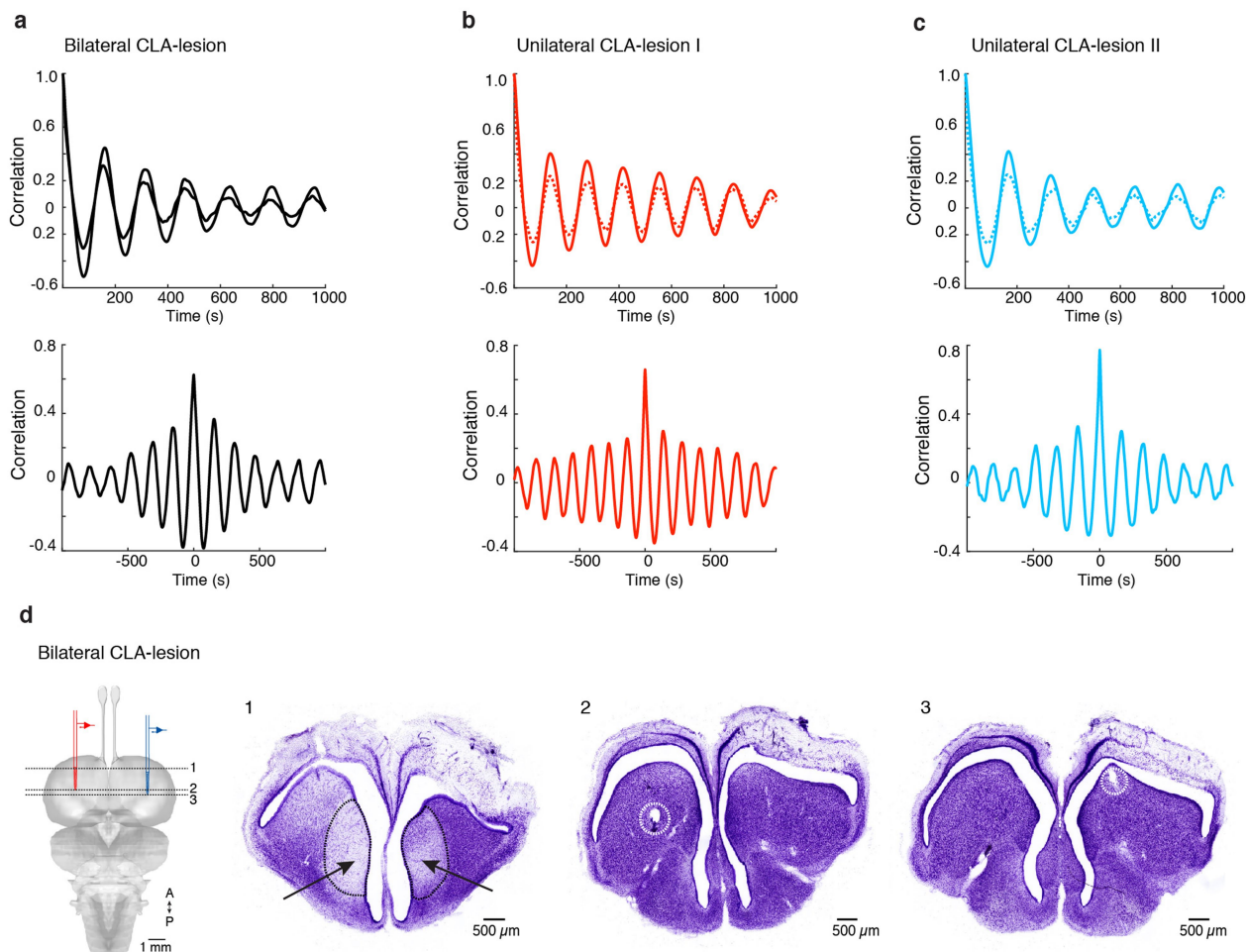
a, Transcriptomic similarity between turtle and lizard clusters, measured as the fraction of single cells that mapped from the turtle pallium dataset to the *Pogona* clusters (Methods). Note that the turtle cell clusters e03–e06 (pallial thickening; PT) map to the lizard cluster 19 (amDVR or claustrum). Turtle data and clusters are from a previous study¹¹. **b**, In situ hybridization in an anterior transverse section, showing expression of the pallial thickening marker gene *crhbp*. Scale bar, 500 μ m. **c**, Architectonics of the lizard claustrum. Right, retrograde labelling of claustrum neurons by rAAV2-retro injected into the aDVR. Left, magnification of the boxed area on the right (in the claustrum). Note the disordered distribution of multipolar neurons. Pink colour shows anti-hippocampal immunostaining. Scale bars, 100 μ m (left); 500 μ m (right). **d**, Architectonics of the turtle claustrum. Right, retrograde labelling of

claustrum neurons by rAAV2-retro injected into the dorso-medial cortex. Left, magnification of the boxed area on the right. Note the arrangement of bipolar neurons within the pallial thickening layer (see also **b** for layering of pallial thickening). Scale bars, 100 μ m (left); 500 μ m (right). **e**, Spontaneous sharp waves recorded simultaneously in the claustrum and the DVR in turtle slice preparation. The red dots in the schematic indicate recording sites. Note sharp wave (LFP) and ripple in the high-pass (HP) band. **f**, Bottom, 295 successive spontaneous ripples (high-pass signal intensity (HPI) > 70 Hz) aligned on the trough of each sharp wave. Top, average of 295 sharp waves aligned on waveform troughs. Grey shading represents s.d. **g**, Representative cross-correlogram of LFP traces recorded simultaneously from the claustrum and the DVR (with claustrum as reference), showing the sharp waves from the DVR trailing those from the claustrum.



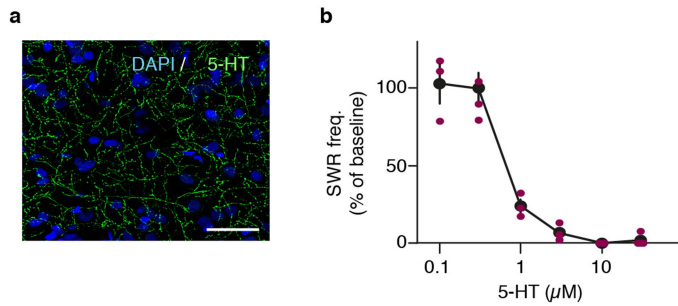
Extended Data Fig. 8 | SWR recordings and stimulation experiments with lizard ex vivo brain preparations. a–f. Experiments in ex vivo brain preparation after cortex removal. **a**, Top, ex vivo brain preparation. Bottom, spontaneous SWRs recorded in the claustrum (<150 Hz) (top trace); HP: 70–150-Hz filtered LFP, showing ripples (bottom trace). **b**, Local pressure injection of 20 μ M TTX into the claustrum and post hoc assessment of injection with Evans blue (transverse section at the bottom, red). **c**, Injection of TTX into the claustrum (shading) silences sharp-wave activity in the claustrum, but also (indirectly) in the DVR. **d**, Analysis of four experiments as in **c**. The filled circles represent mean \pm s.e.m. Claustrum: $*P = 0.029$, $T = 26$ (two-sided Mann–Whitney rank-sum test); DVR: $*P = 0.029$, $T = 26$ (two-sided Mann–Whitney rank-sum test). **e**, Top, average trace and s.d. (shading) from 3,842 sharp waves recorded from the claustrum of an ex vivo forebrain (alignment on trough). Bottom, HPI (>70 Hz) aligned on sharp-wave trough, showing ripple alignment. **f**, Top, simultaneous recordings from ipsilateral claustrum and DVR in an ex vivo preparation. Bottom, cross-correlation between simultaneous recordings in ipsilateral claustrum and DVR, showing that the claustrum precedes the DVR by around 100 ms. **g**, Peristimulus time histogram for multi-unit activity in the cortex, in response to activation of ipsilateral claustrum in an intact ex vivo forebrain. The experiment was carried out in normal ACSF at room

temperature in the presence of 30 μ M serotonin to suppress spontaneous SWRs in the claustrum and 50 μ M carbachol to raise cortex excitability. The claustrum stimulus consisted of a single 50- μ s electrical pulse, delivered with a bipolar electrode. Cortex multi-unit activity was recorded with a glass micropipette. **h**, Change in cortical firing rate (FR) measured in a 200-ms bin after the claustrum stimulus versus a 200-ms bin before the stimulus (as in **g**, on each side of $t = 0$). The control column plots the firing-rate ratio measured in the experiment in **g**, and the GBZ + CGP column plots the results of the same experiment after addition of the GABA receptor antagonists gabazine (GBZ; 5 μ M) and CGP52432 (CGP; 2 μ M); $n = 4$ ex vivo brains from 3 animals each. The control experiment shows that stimulation of the claustrum has an immediate and reliable inhibitory effect on the cortex (#: significantly different from baseline, $P = 0.017$, $t_3 = 4.8$ (two-sided paired t -test)). The stimulation experiment in GABA receptor antagonists shows that stimulation of the claustrum now slightly excites the cortex (**: significantly different from control, $P = 2.0 \times 10^{-3}$, $t_6 = -5.22$ (two-sided Student's t -test)), suggesting that projections from the claustrum both activate and inhibit cortical neurons, probably via direct excitatory projections and indirect inhibitory ones through interneurons (see rodent experiments in a previous study³⁹). Short horizontal lines indicate mean.



Extended Data Fig. 9 | Further analysis of in vivo ibotenic-acid-induced lesion experiments in sleeping *Pogona*. **a**, Autocorrelation (top) and cross-correlation (bottom) of β -band activity in the left and right DVR during sleep in an animal with bilateral claustrum lesions (lesions are shown in **d**). Note that a periodic sleep rhythm (period of around 3 min here) remains after claustrum lesions and therefore does not seem to depend on claustrum integrity. **b**, **c**, Same as **a**, but with unilateral ibotenic-acid-induced lesion in two animals (I and

II). The non-lesioned (sham) side was injected with the same volume of PBS vehicle but without ibotenic acid. Dotted line, sham; solid line, lesion. **d**, Nissl stains (1–3) of transverse sections of the brain of animals with bilateral lesions (shown also in Fig. 4b), at levels indicated in the schematic on the left. Note the claustral lesions (arrows in 1), which are visible as cell-body loss, and the recording sites in left (2) and right (3) DVRs (dotted circles).



Extended Data Fig. 10 | Further data on serotonergic projections to claustrum and their effects on the generation of sharp-wave ripples.

a, Transverse section of claustrum double-labelled with DAPI (blue, nuclei) and serotonin (axonal fibres) antibodies. Note the dense meshwork of serotonergic fibres. Scale bar, 50 μm . **b**, Frequency of spontaneous SWRs in claustrum mini-slices as a function of superfused serotonin concentration. Red circles represent individual experiments (slices). Black points and lines are mean \pm s.e.m.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- ☒ ☐ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☒ ☐ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

in vivo data were collected using Cheetah (Neuralynx) and ex vivo and slice recordings were collected with pClamp 10.5 (Molecular Devices), patchmaster v2x90 (Heka), BrainWave v.4.1. (3Brain), and MC_Rack (Multichannel Systems). For the brain reconstruction we used μ CT scanner and Imaris software (Oxford Instruments), confocal images were taken with Zen 2.1 software (Carl Zeiss).

Data analysis

Custom written code (MatLab 2016a-2017b) was used analyze physiological data. Raw sequencing data were processed using Cell Ranger v3.0 (10xGenomics), and digital gene expression matrices were analyzed in R, using the Seurat v3.0 package. Reconstructed brain images were analyzed with the Voloom 3.0 (micro Dimensions) and Imaris 9.2 (Oxford Instruments).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data and code will be available in full upon request. Accession codes for sequencing/transcriptomic data are provided.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical tests were used to predetermine sample sizes. The amount of brain samples used for each experiment was chosen based on previous experience with this specific type of experiments and commonly used sample sizes in this field of research, taking into account the unusual nature and limited availability of the animal species studied.
Data exclusions	Experiments with off-target placement of electrodes or viral infection were excluded from our analyses. For transcriptomic analyses, only cells were considered for which the number of detected genes was >800 genes/cell (> 1000 for glutamatergic neurons), and the percentage of mitochondrial genes <5%/cell, as detailed in the methods. The exclusion criteria was pre-established.
Replication	We observed similar results which satisfied the same statistical criteria across experiments and we could replicate all our results.
Randomization	Animals were not assigned to groups, and were selected based on weight, health and temperament (for in vivo experiments). Randomization was not relevant for our study.
Blinding	Investigators were not blinded to group allocation during data collection and analysis. Our study was mostly observational in nature with the exception of pharmacological experiments (Fig. 6), comparing the effect of neuromodulators, ibotenic acid, and TTX injections on sleep EEG and SWR production. Measurement of the effect were fully automated, and blinding was thus not relevant for our study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used	Standard, commercially available antibodies were used. Primary antibodies were anti-GFP (A10262, invitrogen, chicken, 1:1000); anti-Hippocalcin (ab24560, abcam, rabbit, 1:1000); anti-ChAT-choline acetyltransferase (AB144P, Merk, goat, 1:100); antiTH-tyrosine hydroxylase (22941, Immunostart, mouse, 1:100 or AB152, Merk, rabbit, 1:200); anti-Histamine (22939, Immunostart, rabbit, 1:100); and anti-Serotonin (MAB352, Merk, rat, 1:100); secondary antibodies were Donkey or Goat anti-rabbit, chicken, goat, mouse, or rat, conjugated with Alexa-488, 568, 647 (A21206, A21208, A11039, A11011, A11057, A11004, A31573, A21247, A31571, all from Invitrogen, all 1:500)
Validation	The manufacturer validated the antibody by Western dot and dot blot. IHC validation was performed in our laboratory, testing various concentrations on lizard tissue.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	Lizards (<i>Pogona vitticeps</i> , either sex, adult (100-400g)) were bred in-house or obtained from external breeders. Turtles (<i>Trachemys scripta elegans</i> or <i>Chrysemys picta</i> , either sex, adult (200-400g)) were obtained from an open-air breeding colony (NASCO Biology, WI, USA), and all species used were housed in our state-of-the-art animal facility.
--------------------	--

Wild animals

This study did not involve wild animals.

Field-collected samples

This study did not involve field-collected samples.

Ethics oversight

All experimental procedures were performed in accordance with German animal welfare guidelines: permit #V54- 19c 20/15-F126/1005 delivered by the Regierungspraesidium Darmstadt, Germany (Dr. E. Simon).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

ATP13A2 deficiency disrupts lysosomal polyamine export

<https://doi.org/10.1038/s41586-020-1968-7>

Received: 21 December 2018

Accepted: 2 December 2019

Published online: 29 January 2020

Sarah van Veen^{1,14}, Shaun Martin^{1,14}, Chris Van den Haute^{2,3}, Veronick Benoy¹, Joseph Lyons⁴, Roeland Vanhoutte⁵, Jan Pascal Kahler⁵, Jean-Paul Decuypere^{1,6,7,8}, Géraldine Gelders², Eric Lambie^{9,10}, Jeffrey Zielich⁶, Johannes V. Swinnen⁶, Wim Annaert⁷, Patrizia Agostinis^{8,11}, Bart Ghesquière¹², Steven Verhelst^{5,13}, Veerle Baekelandt², Jan Eggermont¹ & Peter Vangheluwe^{1*}

ATP13A2 (PARK9) is a late endolysosomal transporter that is genetically implicated in a spectrum of neurodegenerative disorders, including Kufor-Rakeb syndrome—a parkinsonism with dementia¹—and early-onset Parkinson's disease². ATP13A2 offers protection against genetic and environmental risk factors of Parkinson's disease, whereas loss of ATP13A2 compromises lysosomes³. However, the transport function of ATP13A2 in lysosomes remains unclear. Here we establish ATP13A2 as a lysosomal polyamine exporter that shows the highest affinity for spermine among the polyamines examined. Polyamines stimulate the activity of purified ATP13A2, whereas ATP13A2 mutants that are implicated in disease are functionally impaired to a degree that correlates with the disease phenotype. ATP13A2 promotes the cellular uptake of polyamines by endocytosis and transports them into the cytosol, highlighting a role for endolysosomes in the uptake of polyamines into cells. At high concentrations polyamines induce cell toxicity, which is exacerbated by ATP13A2 loss due to lysosomal dysfunction, lysosomal rupture and cathepsin B activation. This phenotype is recapitulated in neurons and nematodes with impaired expression of ATP13A2 or its orthologues. We present defective lysosomal polyamine export as a mechanism for lysosome-dependent cell death that may be implicated in neurodegeneration, and shed light on the molecular identity of the mammalian polyamine transport system.

ATP13A2 is a P5B-ATPase that belongs to the family of P-type ATPases, which couple ATP hydrolysis to substrate transport while transiently forming a catalytic phospho-intermediate⁴. ATP13A2 is generally described as a heavy-metal transporter⁵, but Ca²⁺ (ref. ⁶) and the polyamine spermidine (SPD)^{7,8} have also been proposed as potential substrates. To screen for the transported substrate(s) of ATP13A2, we measured ATPase activity in the presence of various candidate substrates in solubilized microsomal membrane fractions of SH-SY5Y cells that overexpress wild-type human ATP13A2 (hereafter denoted WT-OE) or comparable levels of the catalytically dead mutant ATP13A2(D508N) (containing an aspartic-acid-to-asparagine mutation at position 508; denoted D508N-OE)^{9,10}.

The ATPase activity of wild-type ATP13A2 was significantly stimulated by the polyamines SPD and spermine (SPM) (Fig. 1a), whereas SPM had no effect on the activity of the D508N mutant (Extended Data Fig. 1a). MnCl₂, ZnCl₂, FeCl₃, CaCl₂, diamines, monoamines and amino

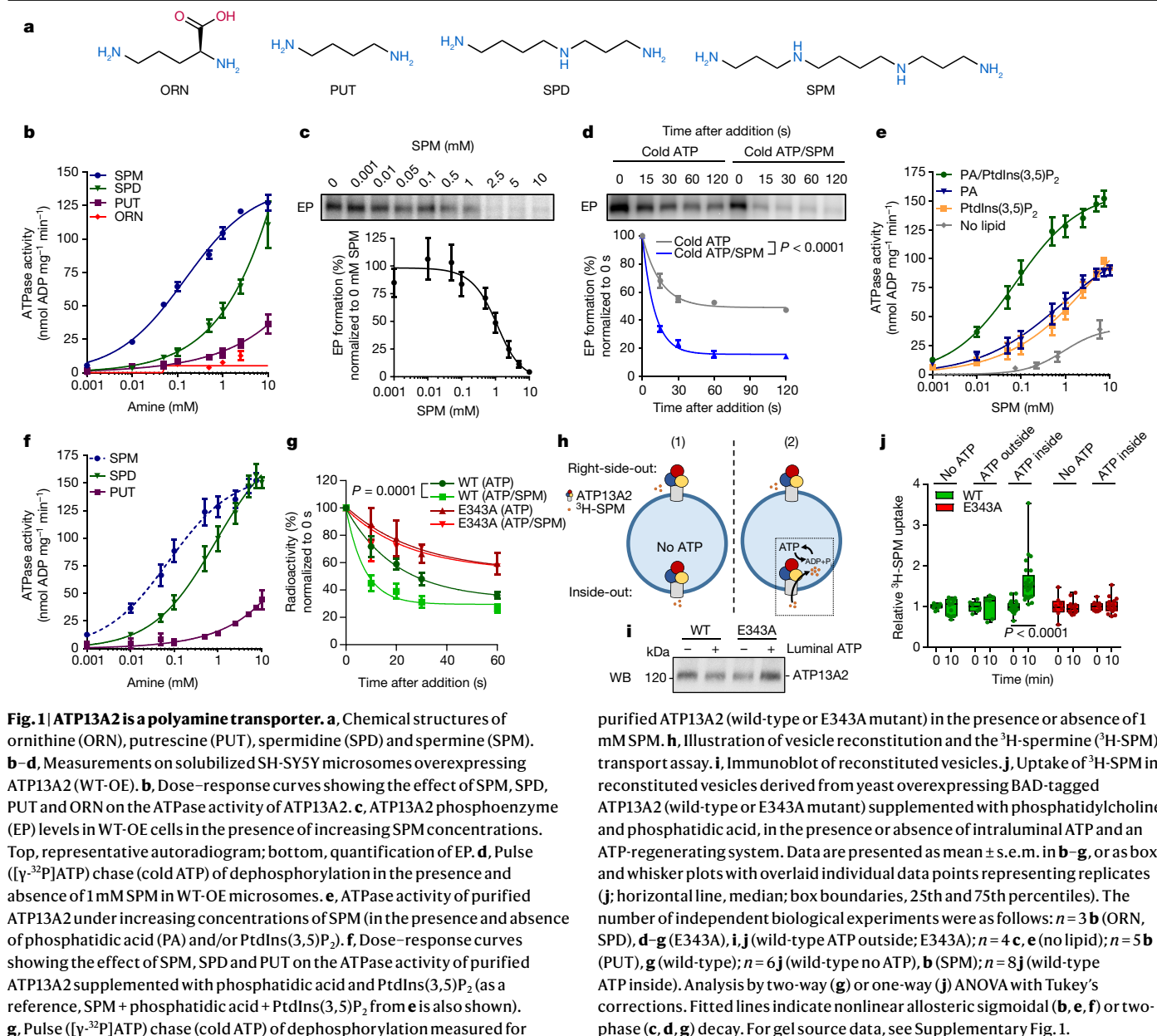
acids exerted no effect on activity (Extended Data Fig. 1a–d), but the polyamines SPM, N¹-acetylspermine and SPD stimulated the ATPase activity of ATP13A2 in a concentration-dependent manner (Fig. 1b, Extended Data Fig. 1e), with the highest apparent affinity observed for SPM (Extended Data Table 1).

The catalytic autophosphorylation and/or dephosphorylation reactions of P-type ATPases occur in response to binding of the transported substrate⁴. ATP13A2 forms a phospho-intermediate on the D508 residue in the absence of SPM supplementation^{9,10}, whereas the addition of SPM leads to a dose-dependent reduction in the levels of the ATP13A2 phosphoenzyme (Fig. 1c), which is not seen with ornithine (Extended Data Fig. 1f). The dephosphorylation rate after a chase with non-radioactive ATP increased in the presence of SPM (Fig. 1d), which further indicates that SPM could be the transported substrate.

SPM also stimulated the ATPase activity of purified human ATP13A2 (for purification, see Extended Data Fig. 2a–e). However, purified

¹Laboratory of Cellular Transport Systems, Department of Cellular and Molecular Medicine, KU Leuven, Leuven, Belgium. ²Laboratory for Neurobiology and Gene Therapy, Department of Neurosciences, KU Leuven, Leuven, Belgium. ³Leuven Viral Vector Core, KU Leuven, Leuven, Belgium. ⁴Department of Molecular Biology and Genetics – DANDRITE, Aarhus, Denmark.

⁵Laboratory of Chemical Biology, Department of Cellular and Molecular Medicine, KU Leuven, Leuven, Belgium. ⁶Laboratory of Lipid Metabolism and Cancer, Department of Oncology, LKI – Leuven Cancer Institute, KU Leuven, Leuven, Belgium. ⁷Laboratory of Membrane Trafficking (VIB-KU Leuven Center for Brain and Disease Research), Department of Neurosciences, KU Leuven, Leuven, Belgium. ⁸Laboratory of Cell Death Research and Therapy, Department of Cellular and Molecular Medicine, KU Leuven, Leuven, Belgium. ⁹Cell and Developmental Biology, Department of Biology II, Ludwig-Maximilians-Universität Munich, Munich, Germany. ¹⁰Department of Cell and Developmental Biology, University College London, London, UK. ¹¹VIB-KU Leuven Center for Cancer Biology, Department of Oncology, KU Leuven, Leuven, Belgium. ¹²Metabolomics Expertise Center (VIB-KU Leuven Center for Cancer Biology), Department of Oncology, KU Leuven, Leuven, Belgium. ¹³Leibniz Institute for Analytical Sciences ISAS, Dortmund, Germany. ¹⁴These authors contributed equally: Sarah van Veen, Shaun Martin. *e-mail: peter.vangheluwe@kuleuven.be



ATP13A2 presented similar properties to microsomal ATP13A2 only in the presence of the regulatory lipids phosphatidylinositol(3,5)bisphosphate (PtdIns(3,5)P₂) and phosphatidic acid, which bind to the N terminus of ATP13A2^{9–11} (Fig. 1e, f, Extended Data Fig. 2f, Extended Data Table 1). SPM-induced ATPase activity was blocked by orthovanadate, a general P-type ATPase inhibitor (Extended Data Fig. 1g). Finally, we also purified the mutant ATP13A2(E343A), which carries a mutation in the conserved catalytic motif for dephosphorylation (³⁴⁴TGES) (Extended Data Fig. 2g). The E343A mutant underwent autophosphorylation (Fig. 1g) but displayed limited SPM-induced ATPase activity (Extended Data Fig. 2h). Notably, when the phosphoenzyme was chased with cold ATP, SPM clearly stimulated dephosphorylation of purified wild-type ATP13A2, but not of the E343A mutant (Fig. 1g, Extended Data Fig. 1h, 6a).

ATP13A2 is a lysosomal polyamine exporter

Next, we performed transport assays with ^3H -labelled SPM (^3H -SPM) in reconstituted vesicles from solubilized yeast membranes expressing biotin acceptor domain (BAD)-labelled wild-type ATP13A2 or the

mutant ATP13A2(E343A), supplemented with the activating lipid phosphatidic acid. This reconstitution rendered two populations of ATP13A2 proteins that were inserted either right-side-out (ATP-binding domain in the extraluminal space) or inside-out (ATP-binding domain in the lumen) (Fig. 1h, i). Uptake of ^3H -SPM was detected for wild-type ATP13A2—but not for the E343A mutant—only when ATP was present inside the vesicles, together with an ATP-regenerating system (Fig. 1j); that is, when ATP13A2 was positioned inside-out ((2) in Fig. 1h). Extending these insights to the cellular context, in which ATP13A2 is present in the late endolysosomal compartment^{1,12}, ATP13A2 most probably operates as a lysosomal SPM exporter.

Notably, the functionality of ATP13A2 affects the cellular polyamine content. We generated two independent ATP13A2-knockout SH-SY5Y cell lines using CRISPR-Cas9 genome editing (denoted KO; Extended Data Fig. 3a) and, upon analysis by mass spectrometry, demonstrated that the total cellular polyamine content was lower in KO than in control cells (Fig. 2a). Expression of either the wild-type ATP13A2 in the KO background (denoted KO/WT) for rescue, or the D508N mutant (KO/D508N) as a negative control (Extended Data Fig. 3b), resulted in

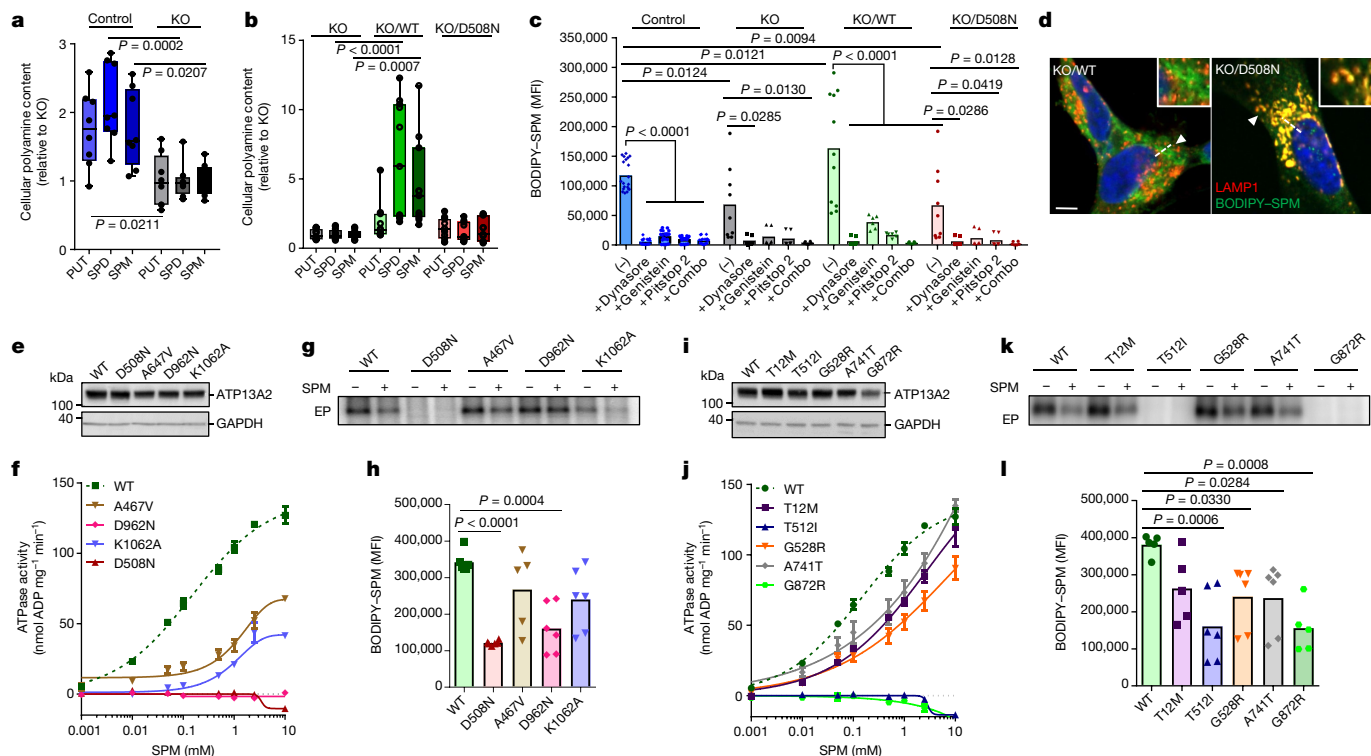


Fig. 2 | ATP13A2 transport affects cellular polyamine uptake, which is impaired by catalytic and disease mutations. a, b, Metabolomics of cellular polyamines in ATP13A2 knockout cells (KO) compared with SH-SY5Y controls (a), or compared with rescue cell lines expressing wild-type ATP13A2 (KO/WT) or ATP13A2(D508N) (KO/D508N) (b). **c,** Uptake of BODIPY-SPM in the presence or absence of the endocytosis inhibitors Dynasore, genistein and/or Pitstop 2, alone or in combination (combo). **d,** Confocal microscopy of BODIPY-SPM distribution (see Methods). Scale bar, 5 μ m. The arrow head shows the region that is expanded in the insets; the dashed line shows the region analysed in Extended Data Fig. 1j. **e, f,** Immunoblotting (e) and SPM-induced ATPase activity (f) of microsomes expressing wild-type ATP13A2, the D508N mutant, or catalytic mutants in M4 (A467V), M6 (D962N) and M8 (K1062A)³⁰. **g,** Representative autoradiogram depicts phosphoenzymes of wild-type ATP13A2 compared with the indicated mutants in the presence or absence of 1 mM SPM. **h,** Uptake of BODIPY-SPM in cells expressing the indicated mutant proteins. **i, j,** Immunoblotting (i) or SPM-induced ATPase activity (j) of SH-SY5Y microsomes overexpressing mutants associated with Kufor-Rakeb syndrome (T512I and G872R) or early-onset Parkinson's disease (T12M, G528R, A741T).

a significantly higher SPD and SPM content in KO/WT cells than in KO or KO/D508N cells (Fig. 2b).

Flow cytometry experiments revealed that ATP13A2 promotes the cellular uptake of BODIPY-labelled polyamines¹³. WT-OE cells took up more BODIPY-SPD and BODIPY-SPM than D508N-OE cells or control cells that expressed firefly luciferase (Fluc; Extended Data Fig. 4a, b). Similarly, KO/WT cells displayed a twofold higher uptake of BODIPY-SPM than ATP13A2 KO and KO/D508N cells, both of which took up less than control cells (Fig. 2c). The comparable uptake of the fluorescein isothiocyanate (FITC)-dextran conjugate in the KO/WT and the KO/D508N cells (Extended Data Fig. 4c) demonstrated that the higher BODIPY-SPM uptake in KO/WT cells compared with KO/D508N cells is not explained by an increased endocytic rate, but depends on the transport activity of ATP13A2. The stimulatory effect of ATP13A2 on endocytosis appears as a transport-independent phenotype¹⁴.

The observation that ATP13A2 transports SPM towards the cytosol (Fig. 1j) and increases the cellular SPM content (Fig. 2a, b) suggests that it could transport endocytosed polyamines into the cytosol. Indeed, endocytosis inhibitors prevented the uptake of FITC-dextran

The SPM dose-response curve from Fig. 1b is shown as a reference in f and j. **k,** Representative autoradiogram depicting phosphoenzymes of wild-type ATP13A2 compared with the disease-related mutants in the presence and absence of 1 mM SPM. **l, m,** BODIPY-SPM uptake in cells containing the indicated mutant proteins. MFI, mean fluorescence intensities; (-), vehicle-treated sample. Data are presented as individual data points (representing replicates) overlaid on box and whisker plots (a, b, horizontal line, median; box boundaries, 25th and 75th percentiles) or mean (c, h, i) or mean \pm s.e.m. (f, j). The number of independent biological experiments were as follows: $n = 3$ d-f, g (D508N, A467V, D962N, K1062A (SPM)), h, j (T12M, T512I, G872R), k (T12M (-), T12M, T512I, G528R, A741T, G872R (SPM)), l; $n = 4$ a-c (KO, KO/WT, KO/DN), g (D508N, A467V, D962N, K1062A (-)), i, j (A741T, G528R), k (wild-type, T512I, G528R, A741T, G872R (-)); $n = 5$ c (control), g (wild-type (-)); $n = 6$ g (wild-type (SPM)), k (wild-type (SPM)). Analysis was performed using one-way ANOVA with Tukey's (a-c) or Dunnett's (h-i) corrections. Fitted lines indicate nonlinear allosteric sigmoidal (f (D962N, D508N), j) or one-phase (f (A467V, K1062A)) association. For gel source data see Supplementary Fig. 1.

(Extended Data Fig. 4c) and blocked that of BODIPY-SPM (Fig. 2c). Using confocal microscopy, we confirmed that KO/WT cells had a higher BODIPY-SPM content than KO/D508N cells (Extended Data Fig. 1i). In the KO/D508N cells, BODIPY-SPM mainly colocalized with LAMP1-positive vesicles (Fig. 2d, Extended Data Fig. 1j, k); this is indicative of accumulation in the late endolysosomes, which is not a lysosomotropic effect¹⁵. By contrast, cells with functional ATP13A2 displayed a broader distribution of BODIPY-SPM, and it was more abundant in the cytosol and nucleus; this is consistent with the transport direction from lysosomal lumen to cytosol (Fig. 2d, Extended Data Fig. 1l). By stimulating cellular SPM uptake and transporting SPM into the cytosol, ATP13A2 complements endogenous SPM synthesis, which depends on the enzymes ornithine decarboxylase (ODC) and the SPD and SPM synthases. Consistent with this, KO/WT cells were protected against pharmacological inhibition of ODC and the SPD and SPM synthases, and a lack of ATP13A2 activity sensitized KO and KO/D508N cells to inhibition (Extended Data Fig. 4d-g); this is consistent with the negative genetic interactions between ATP13A2 orthologues and ODC in yeast¹⁶ and *Caenorhabditis elegans*⁸.

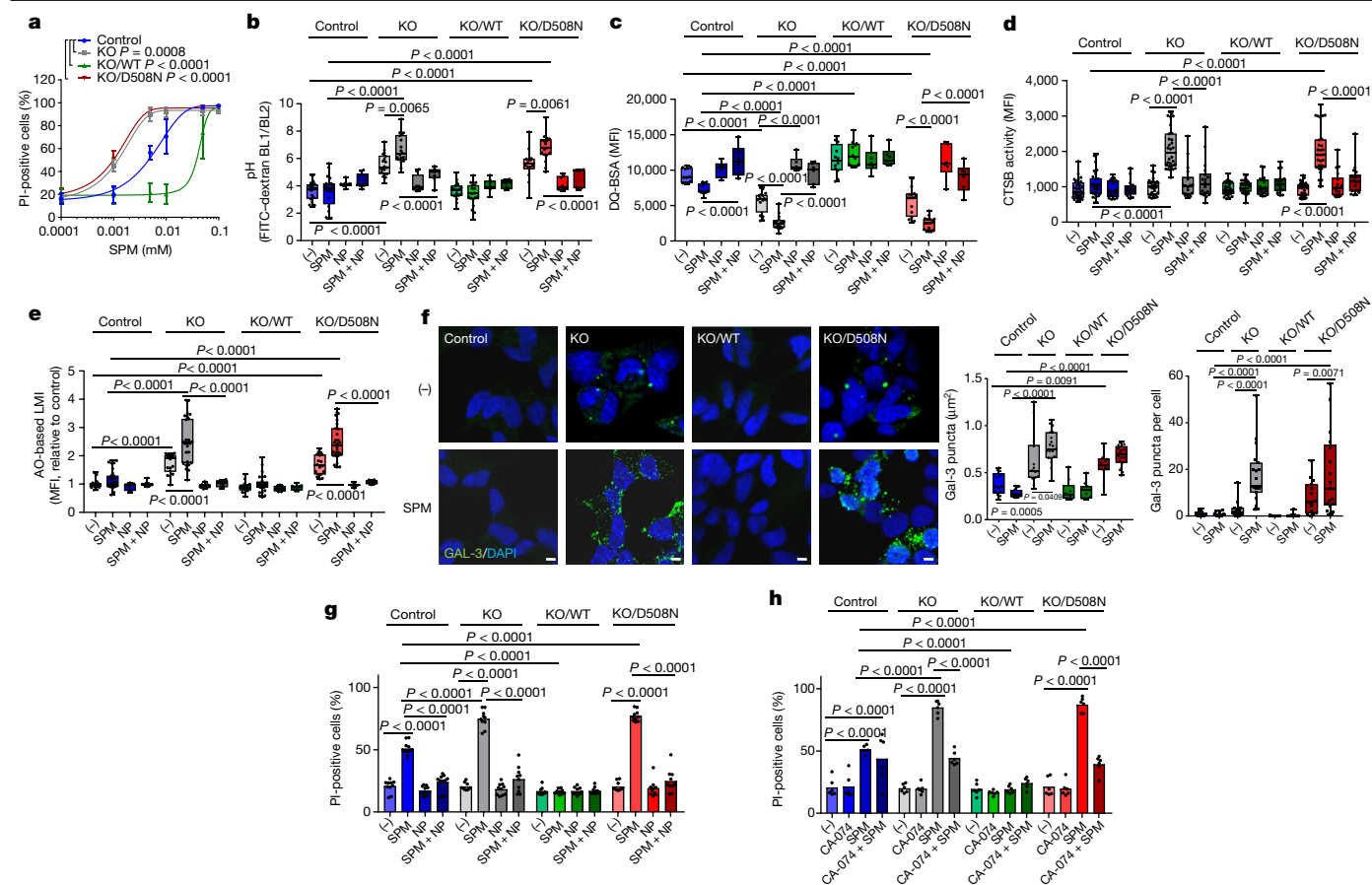


Fig. 3 | ATP13A2 protects against lysosome-dependent SPM toxicity. **a**, SPM-induced cell death (PI, propidium iodide) in SH-SY5Y control, ATP13A2 knockout (KO) and rescue cell lines with wild-type ATP13A2 (KO/WT) or ATP13A2(D508N) (KO/D508N). **b–f**, Effect of SPM (10 μ M, 4 h) and acidic nanoparticles (NP, 100-nm diameter; Extended Data Fig. 8j) on lysosomal functionality. **b**, Lysosomal pH measured by ratiometric FITC–dextran (standard curve, Extended Data Fig. 8i). **c**, **d**, Lysosomal degradation capacity analysed by DQ-BSA (**c**) or cathepsin B (CTSB) (**d**) activity. **e**, **f**, Assessment of lysosomal membrane integrity (LMI) via acridine orange (AO) staining (**e**), or galectin-3 (Gal-3) punctae formation (lysosomal rupture) (**f**). In **f**, the confocal images depict representative images (DAPI staining for nuclei reference; scale

bars, 2.5 μ m). The box and whisker plots depict the size (left) and number (right) of punctae. **g**, **h**, The effect of nanoparticles (**g**) or the cathepsin B inhibitor CA-074 (25 μ M) (**h**) on SPM-induced cytotoxicity. Data are presented as individual data points (representing replicates) overlaid on box and whisker plots (**b–f**, horizontal line, median; box boundaries, 25th and 75th percentiles) or means (**g**, **h**) or mean \pm s.e.m. (**a**). The number of individual biological experiments were as follows: **n** = 3 **a**, **b–d** (NP, NP + SPM), **f–h**; **n** = 4 **e**, (NP, NP + SPM); **n** = 6 **b–d** (–), SPM; **n** = 7 **e** (–), SPM). Analysis was performed using two-way ANOVA with Dunnett's test (**a**) or one-way ANOVA with Tukey's (**b–e**, **g**, **h**) or Sidak's (**f**) test. Fitted lines indicate nonlinear log(inhibitor) versus response (variable slope) (**a**).

Mutations disrupt SPM-induced ATP13A2 activity

We next used mutagenesis techniques to confirm that the SPM-dependent activation of ATP13A2 depends on residues in the predicted substrate-binding site near transmembrane segment M4. We introduced mutations in transmembrane segments M4 (A467V), M6 (D962N) and M8 (K1062A) (Fig. 2e, Extended Data Fig. 5a, b; see Methods for the rationale in choosing these residues). In comparison with the wild type, the three mutants displayed a lower SPM-induced ATPase activity and apparent affinity for SPM (Fig. 2f), as well as a reduced cellular uptake of BODIPY–SPD and BODIPY–SPM (Extended Data Fig. 5c, Fig. 2h), which suggests that these residues contribute to SPM coordination in the membrane region. SPM-induced dephosphorylation was completely abolished in the D962N mutant (Fig. 2g, Extended Data Figs. 5d, 6), indicating that D962 may couple SPM binding to the dephosphorylation reaction.

More than thirty disease-associated mutations have been identified in ATP13A2 (Extended Data Fig. 7). We determined the activity of mainly lysosomal-localized mutants of ATP13A2 arising from point mutations in ATP13A2 that are linked to early-onset Parkinson's disease or Kufor-Rakeb syndrome. The T512I and G872R mutants—associated with Kufor-Rakeb syndrome—did not exhibit ATPase or autophosphorylation

activity and were SPM-insensitive, which is in line with their strongly reduced BODIPY–SPD and BODIPY–SPM uptake (Fig. 2i–l, Extended Data Fig. 5e–g). The T12M, G528R and A741T mutants—associated with early-onset Parkinson's disease—had a less severe effect. Compared with wild type, a reduction in the apparent SPM affinity was observed and phosphoenzyme levels were more (T12M) or less (G528R and A741T) sensitive to SPM. In cells, the uptake of BODIPY–SPD and BODIPY–SPM was less severely impaired with the mutations associated with early-onset Parkinson's disease than those associated with Kufor-Rakeb syndrome (Fig. 2i–l, Extended Data Fig. 5e–g). In summary, ATP13A2-dependent polyamine transport is disturbed in all mutants tested, and the degree of functional impact correlates with the phenotypic differences between early-onset Parkinson's disease and Kufor-Rakeb syndrome; however, the mutation type is not the sole determinant of the clinical phenotype¹⁷.

ATP13A2 protects against polyamine toxicity

We further investigated whether defective lysosomal polyamine export explains the lysosomal phenotype in ATP13A2 KO cells³. SPM and SPD are abundant organic polycations that support cell function, but at

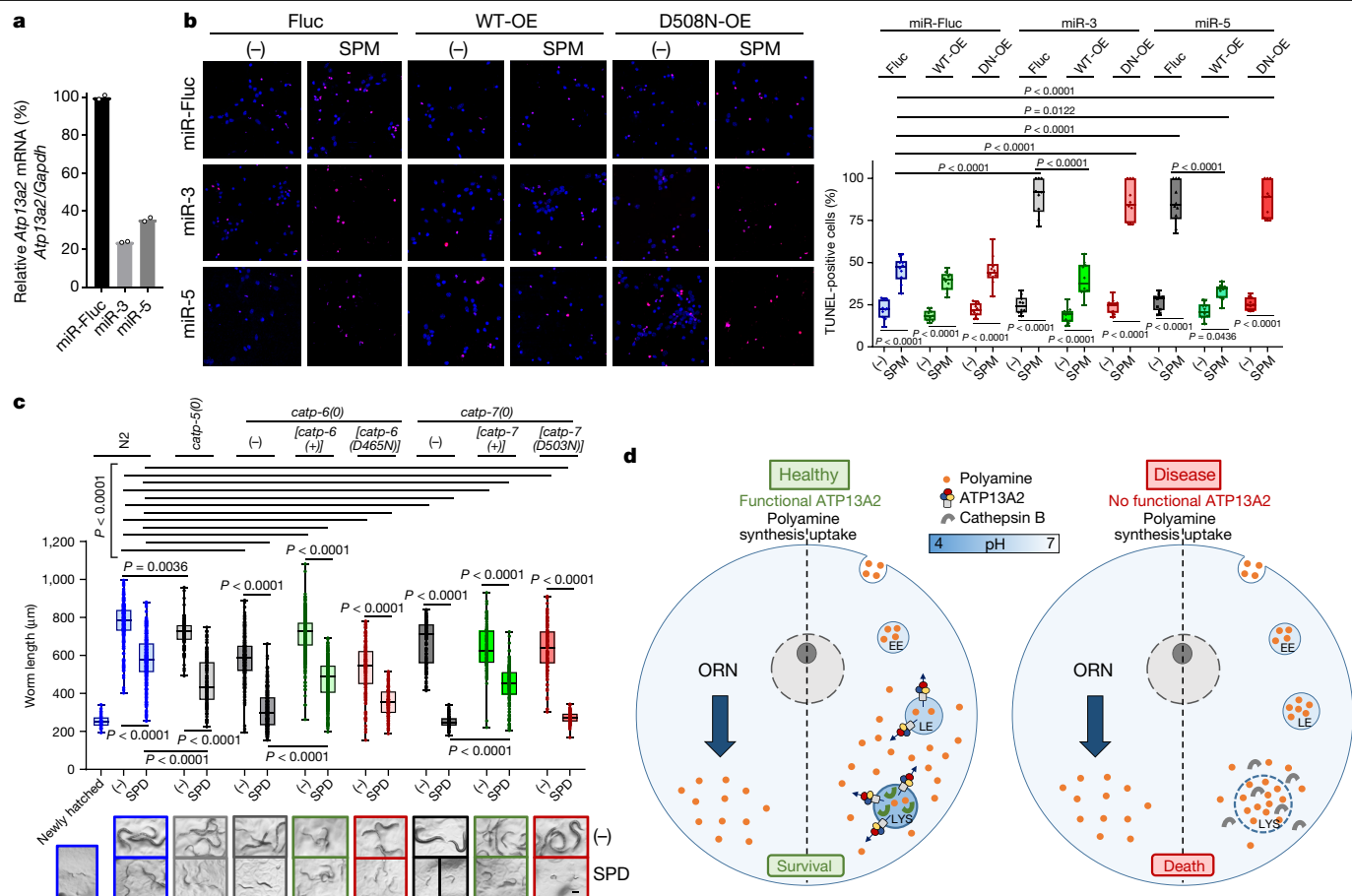


Fig. 4 | Loss of ATP13A2 orthologues exacerbates the toxicity of polyamines in primary neurons and in *C. elegans*. **a**, Lentiviral knockdown of *Atp13a2* in isolated mouse cortical neurons (miR-3 or miR-5) was confirmed by quantitative PCR relative to *Gapdh*. miRNA against firefly luciferase (miR-Fluc) was used as a negative control. **b**, miR-Fluc and *Atp13a2* knockdown neurons were transduced with Fluc, wild-type human ATP13A2 or ATP13A2(D508N). SPM-induced cytotoxicity was assayed via TUNEL staining. Left, representative confocal images depicting TUNEL-positive cells; right, box and whisker plots showing quantification of the TUNEL staining. **c**, The indicated worm strains were assessed for SPD toxicity. Worm lengths were determined as a read-out

high concentrations they become toxic¹⁸. Consistent with this, high levels of SPM or SPD reduced the viability of control cells after 24 hours (Extended Data Fig. 8a, b), which was paralleled by an increase in cell death (Fig. 3a); ornithine and putrescine were not cytotoxic (Extended Data Fig. 8c, d). Notably, loss of ATP13A2 activity exacerbated the toxicity of SPM and SPD (Fig. 3a, Extended Data Fig. 8a, b), which may be a direct consequence of lysosomal polyamine accumulation (Fig. 2d) leading to lysosomal dysfunction. At a time point preceding cell death (4 hours; Extended Data Fig. 8e), lysosomal acidification was compromised in KO and KO/D508N cells, which was aggravated upon SPM exposure (Fig. 3b). This pH-neutralizing effect of SPM was absent in control and KO/WT cells (Fig. 3b).

Lysosomal alkalization may explain the decreased lysosomal degradation potential (Fig. 3c) and cathepsin D activity (Extended Data Fig. 8f) that is observed in KO and KO/D508N cells. However, the activity of cathepsin B increased at toxic SPM levels (Fig. 3d), in line with its higher pH optimum and most likely due to impaired lysosomal membrane integrity¹⁹, which is a driver of lysosome-dependent cell death²⁰. We then confirmed, using an acridine orange-based assay, that lysosomal membrane integrity in ATP13A2 KO and KO/D508N cells was impaired; this was more prominent after SPM challenge,

for toxicity. Scale bar, 100 μ m. Top, quantification; bottom: representative images. **d**, Illustration of the proposed mechanism of endolysosomal polyamine uptake and transfer into the cytosol via ATP13A2. (-), vehicle-treated sample; EE, early endosome; LE, late endosome; LYS, lysosome. Data are presented as mean (**a**) or as box and whisker plots (**b**, **c**, horizontal line, median; box boundaries, 25th and 75th percentiles) with individual data points representing replicates. The number of independent biological replicates were as follows: $n = 2$ (**a**); $n = 3$ (**b**, **c**). Analysis was performed using one-way ANOVA with Tukey's test (**b**, **c**).

a phenotype that was absent in control and KO/WT cells (Fig. 3e). SPM treatment also increased the number and size of endogenous galectin 3 punctae, representative of lysosomal rupture, in KO and KO/D508N cells only (Fig. 3f). Lysosomal rupture was further confirmed by the loss of endolysosomal FITC-dextran punctae and a more diffuse, cytosolic cathepsin B staining in the KO and KO/D508N cells (Extended Data Fig. 8g, h). In line with reported findings²¹, exposure to acidic nanoparticles restored lysosomal pH and functionality (Fig. 3b, c), prevented cathepsin B activation (Fig. 3d) and recovered the intactness of lysosomal membranes (Fig. 3e), ultimately reducing SPM-induced cell death (Fig. 3g). Moreover, pharmacological inhibition of cathepsin B significantly reduced SPM toxicity in the KO and KO/D508N cells (Fig. 3h).

ATP13A2 is protective in higher disease models

Lysosomal polyamine toxicity may be relevant in the context of neurodegeneration, because isolated mouse cortical neurons with miRNA-mediated *Atp13a2* knockdown (Fig. 4a) were more susceptible to SPM-induced cell death than control neurons (Fig. 4b). Notably, the increased sensitivity to SPM observed for neurons in which *Atp13a2* was knocked down was attenuated either by inhibition of cathepsin B

(Extended Data Fig. 9) or by rescue with wild-type human ATP13A2, but not by the D508N mutant (Fig. 4b).

Finally, SPD exposure was found to hamper growth capacity in *C. elegans* N2 (wild-type); this was marked by developmental delay and hence reduced worm length. These effects were exacerbated in worms deficient in ATP13A2 orthologues (*catp-5(O)*, *catp-6(O)* or *catp-7(O)*²²) (Fig. 4c). The phenotype of CATP-6- and CATP-7-deficient strains was rescued by wild-type CATP-6 or CATP-7, but not by a catalytically dead variant, showing that the transport activity is required (Fig. 4c).

Discussion

Polyamines are physiologically important polycations that are tightly regulated by a complex interplay of import, export, synthesis and degradation¹⁸. As a polyamine transporter that controls the cellular polyamine content, ATP13A2 emerges as a member of the mammalian polyamine transport system¹⁸. Extracellular polyamines most probably bind to plasma membrane heparan sulfate proteoglycans²³ and enter the cell through endocytosis²⁴ before transport into the cytosol by ATP13A2 (Fig. 4d). In addition, the other P5B-ATPases (ATP13A3, ATP13A4 and ATP13A5) that reside in the endosomal system share high sequence similarity in the substrate-binding region¹², and may belong to the mammalian polyamine transport system.

Genetic insights suggest a major role for lysosomal dysfunction in Parkinson's disease, contributing to α -synuclein aggregation and mitochondrial dysfunction²⁵. Here we demonstrate that impaired lysosomal polyamine export represents a lysosome-dependent cell death pathway that may be implicated in ATP13A2-associated neurodegeneration. In addition, defective ATP13A2 leads to a reduction in cellular polyamine content, which may potentiate the disease phenotype because polyamines are scavengers of heavy metals and reactive oxygen species²⁶ and regulate autophagy²⁷. The dual effect of ATP13A2 on both lysosomal and cytosolic polyamine levels may explain the broad phenotype that is associated with its loss of function⁵. Other genes that are related to Parkinson's disease may also affect ATP13A2 functionality, or may be affected by disrupted polyamine homeostasis. Polyamine levels decline with age, whereas polyamine supplementation increases lifespan in several model organisms²⁷. Conversely, defective SPM synthase causes Snyder–Robinson syndrome—a form of X-linked intellectual disability²⁸—and reduced expression of SPD/SPM *N*¹-acetyltransferase 1 has been implicated in Parkinson's disease²⁹. Modulation of polyamine homeostasis may therefore be considered for neuroprotective therapy.

In conclusion, ATP13A2 dysfunction prevents late endolysosomal polyamine export and sensitizes cells to lysosomal disruption by exogenous polyamines.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-1968-7>.

- Ramirez, A. et al. Hereditary parkinsonism with dementia is caused by mutations in ATP13A2, encoding a lysosomal type 5 P-type ATPase. *Nat. Genet.* **38**, 1184–1191 (2006).
- Di Fonzo, A. et al. ATP13A2 missense mutations in juvenile parkinsonism and young onset Parkinson disease. *Neurology* **68**, 1557–1562 (2007).
- Dehay, B. et al. Loss of P-type ATPase ATP13A2/PARK9 function induces general lysosomal deficiency and leads to Parkinson disease neurodegeneration. *Proc. Natl Acad. Sci. USA* **109**, 9611–9616 (2012).
- Palmgren, M. G. & Nissen, P. P-type ATPases. *Annu. Rev. Biophys.* **40**, 243–266 (2011).
- van Veen, S. et al. Cellular function and pathological role of ATP13A2 and related P-type transport ATPases in Parkinson's disease and other neurological disorders. *Front. Mol. Neurosci.* **7**, 48 (2014).
- Narayanawamy, N. et al. A pH-correctable, DNA-based fluorescent reporter for organellar calcium. *Nat. Methods* **16**, 95–102 (2019).
- De La Hera, D. P., Corradi, G. R., Adamo, H. P. & De Tezanos Pinto, F. Parkinson's disease-associated human P_{5B}-ATPase ATP13A2 increases spermidine uptake. *Biochem. J.* **450**, 47–53 (2013).
- Heinick, A. et al. *Caenorhabditis elegans* P_{5B}-type ATPase CATP-5 operates in polyamine transport and is crucial for norspermidine-mediated suppression of RNA interference. *FASEB J.* **24**, 206–217 (2010).
- Holemans, T. et al. A lipid switch unlocks Parkinson's disease-associated ATP13A2. *Proc. Natl Acad. Sci. USA* **112**, 9040–9045, (2015).
- Martin, S., Holemans, T. & Vangheluwe, P. Unlocking ATP13A2/PARK9 activity. *Cell Cycle* **14**, 3341–3342 (2015).
- Martin, S. et al. Protection against mitochondrial and metal toxicity depends on functional lipid binding sites in ATP13A2. *Parkinsons Dis.* **2016**, 9531917 (2016).
- Sørensen, D. M. et al. Parkinson disease related ATP13A2 evolved early in animal evolution. *PLoS ONE* **13**, e0193228 (2018).
- Vanhoutte, R., Kahler, J. P., Martin, S., van Veen, S. & Verhelst, S. H. L. Clickable polyamine derivatives as chemical probes for the polyamine transport system. *ChemBioChem* **19**, 907–911 (2018).
- Demirsoy, S. et al. ATP13A2/PARK9 regulates endo-/lysosomal cargo sorting and proteostasis through a novel PI(3, 5)P₂-mediated scaffolding function. *Hum. Mol. Genet.* **26**, 1656–1669 (2017).
- Seglen, P. O. & Gordon, P. B. Effects of lysosomotropic monoamines, diamines, amino alcohols, and other amino compounds on protein degradation and protein synthesis in isolated rat hepatocytes. *Mol. Pharmacol.* **18**, 468–475 (1980).
- Chatr-aryamontri, A. et al. The BioGRID interaction database: 2017 update. *Nucleic Acids Res.* **45**, D369–D379 (2017).
- Estrada-Cuzcano, A. et al. Loss-of-function mutations in the ATP13A2/PARK9 gene cause complicated hereditary spastic paraplegia (SPG78). *Brain* **140**, 287–305 (2017).
- Pegg, A. E. Functions of polyamines in mammals. *J. Biol. Chem.* **291**, 14904–14912 (2016).
- Qiao, C. et al. Atp13a2 deficiency aggravates astrocyte-mediated neuroinflammation via NLRP3 inflammasome activation. *CNS Neurosci. Ther.* **22**, 451–460 (2016).
- Aits, S. & Jäättelä, M. Lysosomal cell death at a glance. *J. Cell Sci.* **126**, 1905–1912 (2013).
- Bourdenx, M. et al. Nanoparticles restore lysosomal acidification defects: implications for Parkinson and other lysosomal-related diseases. *Autophagy* **12**, 472–483 (2016).
- Zielich, J. et al. Overlapping expression patterns and functions of three paralogous P5B ATPases in *Caenorhabditis elegans*. *PLoS ONE* **13**, e0194451 (2018).
- Belting, M. et al. Glypican-1 is a vehicle for polyamine uptake in mammalian cells: a pivotal role for nitrosothiol-derived nitric oxide. *J. Biol. Chem.* **278**, 47181–47189 (2003).
- Uemura, T., Stringer, D. E., Blohm-Mangone, K. A. & Gerner, E. W. Polyamine transport is mediated by both endocytic and solute carrier transport mechanisms in the gastrointestinal tract. *Am. J. Physiol. Gastrointest. Liver Physiol.* **299**, G517–G522 (2010).
- Klein, C. & Westenberger, A. Genetics of Parkinson's disease. *Cold Spring Harb. Perspect. Med.* **2**, a008888 (2012).
- Ha, H. C. et al. The natural polyamine spermine functions directly as a free radical scavenger. *Proc. Natl Acad. Sci. USA* **95**, 11140–11145 (1998).
- Madeo, F., Eisenberg, T., Pietrocola, F. & Kroemer, G. Spermidine in health and disease. *Science* **359**, eaan2788 (2018).
- Li, C. et al. Spermine synthase deficiency causes lysosomal dysfunction and oxidative stress in models of Snyder–Robinson syndrome. *Nat. Commun.* **8**, 1257 (2017).
- Lewandowski, N. M. et al. Polyamine pathway contributes to the pathogenesis of Parkinson disease. *Proc. Natl Acad. Sci. USA* **107**, 16970–16975 (2010).
- Sørensen, D. M., Buch-Pedersen, M. J. & Palmgren, M. G. Structural divergence between the two subgroups of P5 ATPases. *Biochim. Biophys. Acta* **1797**, 846–855 (2010).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Methods

Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Materials

The following reagents were purchased from Sigma-Aldrich: sodium orthovanadate (S6508), CaCl_2 (C3881), ZnCl_2 (Z0152), MnCl_2 (M3634), FeCl_3 (I57740), SPM (S3256), SPD (S2626), N^1 -acetylspermine trihydrochloride (O1467), N^1 -acetylspermidine hydrochloride (9001535-1), N^6 -acetylspermidine dihydrochloride (A3658), putrescine dihydrochloride (P7505), L-arginine (A5006), L-ornithine monohydrochloride (O2375), histamine (H7250), agmatine sulfate salt (A7127), dopamine hydrochloride (H8502), cadaverine (D22606), yeast nitrogen base without amino acids (Y0626), yeast drop-out mix without uracil (Y1501), glucose (G8720), streptavidin sepharose (GE17-5113-01), thrombin (GE27-0846-01), DMSO (276855), DL- α -difluoromethylornithine (DFMO; D193), Dynasore (D7693), Pitstop 2 (SML1169), 4-methylumbelliferyl heptanoate (MUH, M2514), propidium iodide (P4170), CA-074 (C5732), fluorescein isothiocyanate-dextran (FITC-dextran, 46945), DAPI (D9542), anti-ATP13A2 antibody (A3361), anti-GAPDH antibody (G8795), Resomer RG 503H (719870) and SigmaFast protease inhibitor (S8820). In addition, 18:1 PtdIns(3,5)P₂ (1,2-dioleoyl-*sn*-glycero-3-phospho-(1'-myo-inositol-3',5'-bisphosphate) (ammonium salt); 850154), 18:1 phosphatidic acid (1,2-dioleoyl-*sn*-glycero-3-phosphate (sodium salt); 840875) and egg phosphatidylcholine (840051) were obtained from Avanti Polar Lipids. Bovine serum albumin (BSA; 3854.3) was obtained from C. Roth. Yeast extract (103753.0500) was purchased from VWR, and *n*-dodecyl- β -D-maltopyranoside (DDM; 1758-1350) was purchased from Inalco. We obtained Bio-Beads SM-2 resin (1523920) from Bio-Rad. ³H-SPM (ART 0471) was ordered from ARC. APCHA (*N*-(3-aminopropyl) cyclohexylamine; sc-202715) and 4MCHA (*cis*-4-methylcyclohexylamine; sc-272662) were purchased from Santa Cruz Biotechnology. Genistein (ab120112), anti-galectin-3 antibody (ab2785), anti-LAMP1 antibody (ab24170) and anti-cathepsin B antibody (ab58802) were purchased from Abcam. TRYPLE (12604021), AO (A1372) and DQ-Green BSA (D12050) were ordered from Life Technologies.

HEK-293T cells were purchased from ATCC and certified by ATCC via STR genotype analysis. SH-SY5Y cells were purchased from ATCC and certified by ATCC via STR genotype analysis. SH-SY5Y cells from an in-house collection were authenticated via DNA fingerprinting (Leibniz-Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH).

Preparation of compounds and inhibitors

All polyamines, diamines, monoamines and amino acids were prepared to a final stock concentration of 500 mM (200 mM in the case of SPM) in 0.1 M MOPS-KOH (pH 7.0). DFMO was prepared to a final stock concentration of 500 mM in Milli-Q H₂O. The inhibitors 4MCHA and APCHA were dissolved in DMSO to a final stock concentration of 200 mM. The endocytosis inhibitors Dynasore, genistein and Pitstop 2 were dissolved in DMSO to final concentrations of 50 mM, 25 mM and 25 mM, respectively. The cathepsin B inhibitor CA-074 was dissolved in DMSO to a final concentration of 25 mM.

Generation of SH-SY5Y cell models

SH-SY5Y human neuroblastoma cells were transduced with lentiviral vectors to obtain stable overexpression of firefly luciferase (Fluc) or human ATP13A2 (isoform 2, wild-type (ID: NP_001135445), indicated disease or catalytic mutants) and maintained as described previously^{9–11}. The catalytic mutants A467V on M4, D962N on M6 and K1062A on M8 were generated by mutagenesis. P5-type ATPases were discovered by genome sequence analysis 20 years ago and contain highly conserved

motifs for function and substrate binding³¹. The A467V mutation converts PPALP of the predicted substrate-binding site in transmembrane segment M4 into PPVLP that is present in ATP13A5^{4,30}. Also, neighbouring membrane helices contribute to substrate coordination in P-type ATPases, which often relies on conserved and charged residues, such as D962 in M6 and K1062 in M8 of ATP13A2⁴. Furthermore, mutants associated with Kufor-Rakeb syndrome (T512I^{3,32,33} and G872R³⁴) or early-onset Parkinson's disease (T12M², G528R² and A741T^{35–37}) were generated. All cell lines were produced at varying viral vector titres and assessed for equal expression to wild-type ATP13A2.

For CRISPR-Cas9-mediated knockout of ATP13A2, the lentiviral vector lentiCRISPRv2 (Addgene, 52961)³⁸ was used. First, the Cas9 cassette was transformed into a high-fidelity Cas9 by Gibson assembly with a gBlock gene fragment (Integrated DNA Technologies) of the Cas9 portion encoding a protein product with high-fidelity mutations (N497A/R661A/Q695A/Q926A). This Cas9 variant triggers fewer off-target events while retaining its on-target activity³⁹. A single-guide RNA (sgRNA) targeted to *Atp13a2* was designed taking into account a high on-target efficiency, using sgRNA Designer (<https://portals.broadinstitute.org/gpp/public/analysis-tools/sgRNA-design>)⁴⁰, and a low off-target efficiency, via CRISPR Design (<http://crispr.mit.edu/>). A lentiviral CRISPR-Cas9 high-fidelity expression plasmid was created by inserting fragments that contained an sgRNA sequence of ATP13A2 (forward, 5'-CACCGGTCAGGGTCCCATAACCGGT; reverse, 5'-AAACACCGGTTATGGGACCCTGACC) into the lentiCRISPRv2 vector. The generated CRISPR-Cas9 high-fidelity ATP13A2 plasmid (1,000 ng), the packaging plasmid pCMV- Δ R8.91 (900 ng) and the envelope plasmid pMD2G-VSV-G (Addgene, 12259) (100 ng) were mixed together with 200 μ l of JetPrime buffer and 4 μ l of JetPrime reagent (Polyplus-transfection) for transfection of HEK-293T cells according to the manufacturer's protocol. After 4 h at 37 °C and 5% CO₂, the serum-free medium was replaced with DMEM/F12 (Dulbecco's modified Eagle's medium, Nutrient Mixture F-12) supplemented with 10% fetal calf serum (heat-inactivated). After 48 h, the lentiviral vectors were collected by passing the medium through a 0.45- μ m filter, and 0.5 ml of this medium was used to transduce SH-SY5Y cells supplemented with 8 μ g ml⁻¹ polybrene (Sigma-Aldrich). 24 h after transduction, cells were selected in 3 μ g ml⁻¹ puromycin (Sigma-Aldrich) and passaged three times before single clones were isolated via serial dilution. The resulting cells were examined by quantitative PCR (qPCR) (*ATP13A2* forward, 5'-ACCGGTTATGGGACCCTGAC; *ATP13A2* reverse, 5'-GTGATAGCCGATGACCCTCC) with *HPRT* and *TBP* as internal controls (*HPRT* forward, 5'-TGAGGATTTGAAAGGGTGT; *HPRT* reverse, 5'-ACATCTCGA GCAAGACGTTTC; *TBP* forward, 5'-CGGCTGTTAACTTCGCTTC; *TBP* reverse, 5'-CACACGCCAAGAACAGTGA) and western blotting. For rescue experiments, ATP13A2 knockout cells were stably transduced with lentiviral vectors expressing wild-type ATP13A2 or the D508N mutant in which both cDNAs were modified with synonymous mutations at the sgRNA target site. All cell lines were routinely assessed for mycoplasma and cultured for a maximum of 20 passages.

Membrane fractionation

SH-SY5Y cells were seeded in 15-cm dishes at a density of 6×10^6 cells per plate. The cells were collected 24 h later after trypsinization and brief centrifugation (300g, 5 min). Subcellular fractionation was performed by differential centrifugation, as described previously^{9–11}. The microsomal protein concentration was measured using the bicinchoninic acid assay (Thermo Fisher Scientific, Pierce) according to the manufacturer's instructions.

ATPase assay

The ATPase activity of ATP13A2 was assessed using a commercially available luminescence assay (ADP-Glo Max assay, Promega) that monitors the production of ADP via luciferase activity. Substrate screen was designed to include candidates previously postulated in

Article

the literature^{6–8,41–45}. The reactions were performed for 30 min (37 °C) in a final volume of 25 µl. The assay reaction mixture contained 50 mM MOPS-KOH (pH 7), 100 mM KCl, 11 mM MgCl₂, 1 mM DTT, 195 µM DDM, various concentrations of the indicated compound, and either microsomes (5 µg) collected from SH-SY5Y cells overexpressing ATP13A2 (wild-type or mutants) or purified ATP13A2 (0.3–0.5 µg). When purified ATP13A2 was used, we included 125 µM phosphatidic acid, 125 µM PtdIns(3,5)P₂ and 19.5 µM DDM in the reaction buffer. The assay was started by the addition of 5 mM ATP and was terminated by adding 25 µl of ADP-Glo Reagent. The 96-well plate was then incubated for 40 min at room temperature, followed by the addition of 50 µl of ADP-Glo Max Detection Reagent. After 40 min, luminescence was detected using a FlexStation 3.0 system (Molecular Devices). Dose–response curves and n , K_m and V_{max} values were calculated using GraphPad Prism Software (GraphPad Software).

Autophosphorylation assay

The autophosphorylation activity of ATP13A2 on the conserved D508 residue was measured as described previously^{9,12,17}. In brief, microsomes (20 µg) or purified ATP13A2 (1 µg) were incubated with radioactively labelled ATP in the presence of the indicated SPM or ornithine concentrations, and after 1 min the reaction was stopped. In the case of purified ATP13A2, 125 µM phosphatidic acid and 125 µM PtdIns(3,5)P₂ were included in the reaction mixture. To determine the sensitivity of the ATP13A2 phosphoenzyme to ATP or a combination of ATP and SPM, 30 s after adding ³²P-ATP samples were incubated with non-radioactive ATP (5 mM in experiments using microsomes, 1 mM for purified ATP13A2) and SPM (1 mM) before the reaction was stopped at the indicated time points. The incorporation of ³²P was visualized after SDS–PAGE under acidic conditions and subsequently detected by autoradiography (microsomes) or liquid scintillation counting (purified ATP13A2) (Liquid Scintillation Analyzer TRI-CARB 2900TR).

Transformation and overexpression of ATP13A2 in yeast

The *S. cerevisiae* W303-1B/Gal4-ΔPep4 strain (leu2-3, his3-11,15, trp1-1::TRP1-GAL10-GAL4, ura3-1, ade2-1, canr, cir⁺, ΔPep4 MAT; a gift from R. Lopez Marques) was transformed according to the lithium acetate/single-stranded carrier DNA/polyethylene glycol method⁴⁶, with the pYedP60 vector containing a yeast codon-optimized version of human ATP13A2 variant 2 (wild-type or the catalytically dead E343A mutant) followed by a thrombin cleavage site and a C-terminal BAD tag^{47,48}. The transformation mixture was grown for 48 h at 30 °C on minimal medium agar plates lacking uracil (0.54% yeast nitrogen base without amino acids, 0.12% yeast drop-out mix without uracil, 2% glucose and 2% agar) to select yeast colonies that acquired the plasmid. These colonies were then cultured in 20 ml of MM-Ura medium (0.67% yeast nitrogen base without amino acids, 0.19% yeast drop-out mix without uracil, and 2% glucose) and grown for 24 h at 28 °C and 200 rpm. The MM-Ura yeast pre-culture was used to inoculate 100 ml of MM-Ura medium to a final OD₆₀₀ of 0.2, followed by a 12-h incubation period (28 °C and 200 rpm). The second pre-culture was inoculated into 4.5 l of YPGE2X medium (2% yeast extract, 2% bactopectone, 1% glucose and 2.7% ethanol) to a final OD₆₀₀ of 0.05 and grown for 36 h (28 °C and 175 rpm). ATP13A2-BAD expression was induced with 2% galactose, followed by a second galactose induction 12 h later. After another 12 h, the pellet was collected (1,000g, 10 min, 4 °C).

Yeast membrane preparation

Yeast cells were broken with glass beads using a BeadBeater (BioSpec products). The lysis buffer contained 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.6 M sorbitol, 1 mM phenylmethylsulfonyl fluoride and SigmaFast protease inhibitor. To remove cell debris and nuclei, the crude extract was centrifuged at 2,000g for 20 min (4 °C). The supernatant (S1) was centrifuged at 20,000g for 20 min (4 °C) to pellet the heavy membrane fraction (P2), and the resulting supernatant (S2) was further

centrifuged at 200,000g for 1 h (4 °C). The resulting pellet (that is, the light membrane fraction, P3) was resuspended in 20 mM HEPES-Tris (pH 7.4), 0.3 M sucrose and 0.1 mM CaCl₂. The total protein concentration was determined using a Bradford assay (B6916, Sigma-Aldrich).

Purification of ATP13A2–BAD

The method used to purify ATP13A2 was based on the purification of overexpressed Drs2p from yeast membranes⁴⁷. Yeast P3 membranes were diluted to 5 mg of total protein per ml in SSR buffer (50 mM MOPS-KOH (pH 7), 100 mM KCl, 20% glycerol, 5 mM MgCl₂, 1 mM DTT, and SigmaFast protease inhibitor cocktail) and solubilized using DDM, with a detergent-to-protein ratio of 1:1. The samples were stirred on ice for 30 min, followed by centrifugation (100,000g, 1 h, 4 °C) to pellet non-solubilized membranes. The solubilized material was incubated with streptavidin beads for 4 h (4 °C) to enable binding of BAD-tagged ATP13A2 to the resin. To eliminate unbound material, the resin was washed four times with three resin volumes of SSR buffer supplemented with 0.5 mg ml^{−1} DDM. Subsequent cleavage by thrombin (0.0625 U per mg total protein) enabled the release of ATP13A2 from the beads by overnight incubation at 4 °C. Finally, a Vivaspin Turbo 4 concentrator (100 kDa molecular weight cut-off, Sartorius) was used to concentrate the sample. The protein concentration was determined using a Bradford assay. The quality of the purification was evaluated via SDS–PAGE followed by Coomassie staining or immunoblotting, as described previously^{9–11}. Furthermore, the purified ATP13A2 sample was analysed by linear mode MALDI–TOF MS (matrix-assisted laser desorption time-of-flight mass spectrometry; Applied Biosystems 4800 Proteomics Analyzer) in the presence of α-cyano-4-hydroxycinnamic acid as matrix and after C4 omix (Agilent) pipette tip purification.

Reconstitution of yeast membranes

To reconstitute yeast membranes, we followed a similar strategy as described before⁴⁹. P3 membranes from the yeast membrane preparation expressing the ATP13A2–BAD construct were solubilized in buffer T (10 mM Tris-HCl (pH 7.4) and 1 mM EDTA) supplemented with 1.4% DDM. After removing the insoluble fraction by ultracentrifugation (30 min, 200,000g), the detergent extract was supplemented with 4.5 mM egg phosphatidylcholine and 0.5 mM 18:1 phosphatidic acid (in buffer T containing 0.7% DDM). The extract was then treated with Bio-Beads to remove the DDM and reconstitute proteoliposomes (that is, ‘no ATP inside’ condition). To generate proteoliposomes that contained intraluminal ATP (that is, ‘ATP inside’ condition), we added 5 mM ATP and an ATP-regenerating system before incubation with the Bio-Beads. Finally, the vesicles were recovered by ultracentrifugation (1 h, 200,000g) and resuspended in buffer T. The protein concentration was determined using a Bradford assay.

Transport assay using reconstituted vesicles

Uptake of ³H-SPM into freshly prepared vesicles was measured within 60 min. The above-described vesicles (‘no ATP inside’ or ‘ATP inside’) were diluted to 1 µg µl^{−1} in buffer T. The reactions were performed for 10 min (30 °C) in a final volume of 1 ml. The assay reaction mixture contained 50 mM MOPS, 100 mM KCl, 11 mM MgCl₂, 1 mM DTT and reconstituted vesicles (45 µg). The reaction was started by adding 1 mM ³H-SPM. For the condition ‘no ATP inside’, 5 mM ATP and an ATP-regenerating system were added after Bio-Bead treatment, before the addition of ³H-SPM. The reaction was stopped by filtering the samples through Millipore filters (0.45 µm). After washing of the filters with assay buffer, radioactivity retained on the filters was counted using a liquid scintillation counter (Liquid Scintillation Analyzer TRI-CARB 2900TR).

Cellular transport assay and endocytosis assessment

BODIPY–SPD and BODIPY–SPM were synthesized as described previously (compounds 14 and 15, respectively)¹³. The cells were seeded in 12-well plates (1.0 × 10⁵ cells per well) and the next day the cells were

incubated with 5 μM BODIPY-SPM or BODIPY-SPD for 2 h before collection. To assess endocytosis, the cells were pre-treated (30 min) with the endocytosis inhibitors Dynasore (100 μM), genistein (50 μM) and/or Pitstop 2 (50 μM) before the addition of either 20 $\mu\text{g ml}^{-1}$ FITC-dextran or BODIPY-SPM (2 h and 37 °C). The cells were then collected (300g, 5 min), washed and resuspended in PBS containing 1% BSA. Finally, an Attune Nxt (Thermo Fisher Scientific) flow cytometer was used to record the mean fluorescence intensities (MFI) of 10,000 events per treatment.

Metabolomics

Cells were grown into a 6-well plate and extracted as described previously⁵⁰. In brief, the medium was removed and cells were washed with a 0.9% NaCl solution. The washing solution was removed and 150 μl of a 6% trichloroacetic acid (Sigma) was added for the extraction. Using a cell scraper, the full extract was transferred into an eppendorf and incubated for 30 min on ice. Insoluble material, such as precipitated proteins, was removed by centrifugation for 20 min at 20,000g at 4 °C. To 100 μl of the supernatant, 900 μl of a 100 mM sodium carbonate buffer (pH 9.0) was added. Next, 25 μl of isobutyl chloroformate (Sigma) was added and the mixture was incubated for 30 min at 35 °C. 800 μl of the reaction mixture was transferred to a 2 ml eppendorf tube and 1 ml of diethylether (Sigma) was added. The mixture was vortexed vigorously and placed for 15 min at 25 °C. 900 μl of the upper phase was transferred into an eppendorf and dried using a vacuum centrifuge. Finally, the dried extract was dissolved in 125 μl of a 50% acetonitrile (LC-MS grade, Merck) solution in water containing 0.2% acetic acid.

15 μl of the extract was loaded onto a Thermo Scientific Liquid Chromatography QQQ (Quantiva, Thermo Fisher Scientific) equipped with an ACQUITY UPLC BEH C18 (1.7 μm , 2.1 \times 100 mm) column from WATERS. Solvent A consisted of ultrapure H₂O with 0.2% acetic acid while solvent B was acetonitrile (Merck) with 0.2% acetic acid; all solvents used were LC-MS grade. Flow rate remained constant at 250 $\mu\text{l min}^{-1}$, and the column temperature remained constant at 30 °C. A gradient for the separation of modified polyamines was applied as follows: from 0 to 2 min 20% B, from 2 to 10 min a linear increase to 85% B was carried out and 85% B was maintained until 17 min. At 18 min the gradient returned to 20% B. The method stopped at 22 min. The mass spectrometer operated in positive ion mode (3,500 V); the source settings were as follows: sheath gas at 50, aux gas at 10, the ion transfer tube was heated at 325 °C and the vaporizer temperature was set at 350 °C. The mass spectrometer operated in multiple reaction monitoring mode and used the following transitions: putrescine (parent m/z at 289.2 \rightarrow fragment 215.2, collision energy at 10.25 V), SPD (parent m/z at 446.4 \rightarrow fragment 198.2, collision energy at 23.6 V) and SPM (parent m/z at 603.4 \rightarrow fragment 455.1, collision energy at 19.91 V). Peak area was integrated using the XCalibur Quan tool (version 4.2.28.14, Thermo Fisher Scientific).

Preparation and characterization of acidic nanoparticles

Acidic nanoparticles were prepared as previously described²¹. In brief, 31 mg of Resomer RG 503H (lactide to glycolide ratio 50:50, molecular weight 24–38 kDa) was dissolved in 3.1 ml of tetrahydrofuran and subsequently 200 μl of this solution was added to 20 ml of ultrapure water under sonication. The suspension was then concentrated using the rotary evaporator to a final volume of approximately 12 ml, resulting in a concentration of 0.167 mg ml^{-1} . The size distribution of prepared nanoparticles was measured using a Wyatt DynaPro DLS plate reader (Wyatt), using an 830-nm laser in a flat-bottom 384-well plate (Greiner) at 25 °C and 10 measurements were averaged per experiment.

MUH cytotoxicity assay

SH-SY5Y cells were seeded in 96-well plates (1×10^4 cells per well) and allowed to adhere overnight. The cells were subsequently treated with increasing doses of the indicated compounds for 24–48 h. After

exposure, the cells were washed with PBS and stained with 300 $\mu\text{g ml}^{-1}$ MUH (prepared in DMSO, dissolved in PBS) for 30 min at 37 °C. Cytotoxicity was read using a FlexStation 3.0 multi-well plate reader (Molecular Devices; excitation 360 nm, emission 460 nm, cut-off 455 nm). Data were expressed relative to control.

Propidium iodide exclusion assay

Cells were seeded in 12-well plates (1.0×10^5 cells per well) and the next day the cells were treated with increasing doses of SPM, alone or in combination with the cathepsin B inhibitor CA-074 (25 μM , 1 h pre-incubation) or acidic nanoparticles (180 ng ml^{-1} , 1 h pre-incubation), and incubated for 24 h at 37 °C. Thereafter, the cells were collected following trypsinization and a brief centrifugation (300g; 5 min), washed with PBS and stained with 1 $\mu\text{g ml}^{-1}$ propidium iodide (in PBS containing 1% BSA). An Attune Nxt (Thermo Fisher Scientific) flow cytometer was used to determine the proportion of propidium iodide-positive cells (10,000 events per treatment).

FITC-dextran-based lysosomal pH

The protocol was adapted from refs.^{51,52}. SH-SY5Y cells were seeded in 12-well plates (1.0×10^5 cells per well) and allowed to adhere overnight. Cells were exposed to 50 $\mu\text{g ml}^{-1}$ FITC-dextran for 72 h. Samples were then washed and placed in fresh medium for 2 h before treatment with SPM (10 μM) alone or in combination with acidic nanoparticles (180 ng ml^{-1} , 1 h pre-incubation) for a further 4 h. Samples were then collected by centrifugation (300g, 5 min) and washed in PBS. Cells were finally resuspended in 500 μl of PBS containing 1% BSA and FITC dual emission was assessed by flow cytometry (excitation 488 nm, emission 530 nm (BL1) and 600 nm (BL2)) of 10,000 events per condition using an Attune NXT flow cytometer (Thermo Fisher Scientific). The emission ratio (BL1/BL2) of all samples were compared to a standard curve, whereby signals were obtained from untreated cells resuspended in monensin (100 μM) containing Britton Robinson buffer with increasing pH (3.0–8.0).

Lysosomal degradative capacity

Cell lines were seeded in 12-well plates (1.0×10^5 cells per well) and the next day the cells were pre-treated with SPM (10 μM) for 1 h at 37 °C. For samples requiring exposure to acidic nanoparticles, cells were treated with 180 ng ml^{-1} 1 h before the addition of SPM. Subsequently, 5 $\mu\text{g ml}^{-1}$ DQ-Green BSA was added to the cells for a further 3 h (37 °C). Finally, the cells were collected (300g, 5 min), and the MFI of 10,000 events were assessed using an Attune Nxt (Thermo Fisher Scientific) flow cytometer.

Cathepsin activity assays

SH-SY5Y cells were seeded in 10-cm plates (2×10^6 cells per plate) and allowed to adhere overnight before treatment with 10 μM SPM for 4 h at 37 °C. For samples requiring the addition of acidic nanoparticles, 180 ng ml^{-1} nanoparticles were added 1 h before the addition of SPM. Next, the samples were collected using TRYPLE and a brief centrifugation (300g, 5 min). The activities of cathepsin B (ab65300) and cathepsin D (ab65302) were assessed using commercially available kits (Abcam) according to the manufacturer's instructions. MFI values were acquired using a FlexStation 3.0 multi-well plate reader (Molecular Devices).

Lysosomal membrane integrity

SH-SY5Y cells were seeded in 12-well plates (1.0×10^5 cells per well) and the next day the cells were incubated with 5 $\mu\text{g ml}^{-1}$ acridine orange (dissolved in medium) for 15 min at 37 °C. Thereafter, the medium was discarded, the cells were washed, and fresh medium was added. For samples requiring acidic nanoparticles, cells were treated with 180 ng ml^{-1} 1 h before the addition of SPM. The cells were then treated with 10 μM SPM for 4 h at 37 °C. Finally, the cells were collected and resuspended in PBS containing 1% BSA. The MFI of 10,000 events was

captured using an Attune Nxt (Thermo Fisher Scientific) flow cytometer.

BODIPY-SPM localization and lysosomal rupture analysis

For all immunofluorescent stainings, SH-SY5Y cell lines were seeded in 12-well plates (0.25×10^5 cells per well on coverslips) and the next day the cells were incubated with the indicated compounds for 4 h at 37 °C. For BODIPY-SPM analysis, cells were pulsed with 5 μ M BODIPY-SPM for 15 min, washed and chased in fresh medium for a further 105 min at 37 °C. After treatment, the cells were washed twice in PBS, fixed in 4% paraformaldehyde for 30 min (37 °C), washed in PBS and stored at 4 °C. For immunofluorescence staining, the cells were washed in PBS containing 0.5% Tween 20 (PBS-T), permeabilized in PBS containing 0.1% Triton X-100 (30 min) and blocked first in 0.1 M glycine (1 h) and then in PBS-T containing 1% fetal calf serum and 10% BSA (30 min). Gal-3-, LAMP1- and cathepsin B-specific antibodies were used at a dilution of 1:100–200 (in PBS-T containing 1% BSA) overnight at 4 °C. Subsequently, the samples were washed and incubated with Alexa Fluor secondary antibody (1:1,000; 30 min). To assess FITC-dextran release, cells were loaded with 50 μ g ml⁻¹ FITC-dextran for 72 h, cells were then washed and placed in fresh medium for 1 h before the addition of SPM (10 μ M). To visualize the nucleus, all samples were stained with DAPI (200 ng ml⁻¹, 15 min). After staining, the samples were fixed to slides, and images were acquired using an LSM780 or LSM880 confocal microscope (Zeiss). For the acquisition of BODIPY-SPM, images were taken with equal settings to confirm uptake potential of ATP13A2 (Extended Data Fig. 1i). To assess the intracellular distribution of BODIPY-SPM (Fig. 2d, Extended Data Fig. 1j–l), microscope settings were optimized per cell type to enable a comparable assessment of BODIPY-SPM localization in KO/WT and KO/D508N cells.

Neuron isolation

Primary cortical neurons were derived from FVB/N mice embryos at embryonic day (E)16. Pregnant mice were euthanized on gestation day 16 by cervical dislocation. The brains of E16 mouse pups were collected and placed in a dish containing calcium- and magnesium-free Hanks' Balanced Salt Solution (HBSS, Life Technologies, 14180-046) on ice. Both cerebral hemispheres were separated from the cerebellum. Meninges were removed from the cerebral hemispheres and the brain cortices dissected. Brain cortices were collected, washed twice and digested with 0.05% trypsin (Life Technologies, 25300-054, 10 min at 37 °C). The trypsin reaction was terminated by the addition of 7 ml HBSS and 1 ml of horse serum. Cells were separated by pipetting and filtration through a cell strainer (40 μ m, Falcon, 352340). Cells were centrifuged at 1,000 rpm for 5 min (4 °C), the supernatant discarded and the pellet suspended in 5 ml Dulbecco's modified Eagle's medium (DMEM; Sigma-Aldrich, D6546) containing GlutaMAX (Life Technologies, 31966-021), 5% horse serum (Life Technologies, 26050-088) and 20 mM glucose (Sigma-Aldrich, 8270). Primary cortical neurons were plated in 12-well plates, on coverslips coated with poly-D-lysine (Sigma-Aldrich, P6407). After an overnight incubation, cell medium was exchanged for Neurobasal medium (Life Technologies, 21103-049) supplemented with 2 mM L-glutamine (Life Technologies, 25030-24) and 2% B27 (Life Technologies, 17504-044).

Atp13a2 knockdown and rescue in isolated cortical neurons

For knockdown, microRNA (miR)-based short-hairpin lentiviral vectors were generated as described⁵³. The two most potent miRs against mouse *Atp13a2* (mouse miR-3, CCACGCCGAAACACTCGTTATA and mouse miR-5, CGCCGAAACACTCGTTATAGAA) were used to induce knockdown in mouse primary neurons. A miR targeting Fluc was used as a control (miR-Fluc, ACGCTGAGTACTTCGAAATGTC). Primary neurons were transduced 4 days after isolation, 72 h before experimentation. For rescue of ATP13A2 expression in knockdown conditions, neurons were subjected to a second round of transductions with either human

wild-type or the D508N variant of ATP13A2, 24 h after the addition of the miR. Fluc was used as an overexpression control. At day 7 post-isolation, cells were treated with 10 μ M SPM for 24 h. To test the contribution of cathepsin B to SPM-induced cell death, neurons were pre-treated (30 min) with 10 μ M CA-074. Knockdown efficiency was validated with qPCR on mRNA levels 72 h after transduction using the following primers (*Atp13a2* forward, CATGGCCCTCTACAGCCTGA; *Atp13a2* reverse, CTCATGAGCACCGCAACCGT) with *Gapdh* as internal control (forward, TGTGTCCGTCGTGGATCTGA; reverse, CCTGCTTACCACCTTCTTGA). All mouse primary neuron experiments were carried out in accordance with the European Communities Council Directive of November 24 1986 (86/609/EEC) and approved by the Bioethical Committee of the KU Leuven (Belgium) (ECD project P185-2014).

TUNEL staining

TUNEL staining was assessed according to the manufacturer's protocol for the Click-iT Plus TUNEL assay (Thermo Fisher Scientific, C10617). DAPI was used as a nuclear counterstain and images were acquired using a LSM780 confocal microscope.

Caenorhabditis elegans

During routine culture, nematodes were grown on NGM (nematode growth medium) using *Escherichia coli* strain AMA1004 as food source^{54,55}. In order to reduce the likelihood that bacterial catabolism and divalent cations would interfere with SPD assays, these were done using DCDA (divalent-cation-depleted agar) plates, which do not permit bacterial growth. This medium contains 2% agar (Carl Roth 5210) that has been washed with 50 mM EDTA, followed by multiple washes with reverse osmosis-purified water; 50 mM HEPES pH 7.4; and 20 μ g ml⁻¹ kanamycin. SPD trihydrochloride (1 M stock solution; Sigma S2501) was added to a final concentration of 5 mM after micro-waving to melt the agar. Assays were performed in 35-mm plates that contained 1 ml of DCDA. Approximately 2 μ l of *E. coli* was transferred to each assay plate from the lawn of a seeded NGM plate. In each case, multiple adult hermaphrodites were added to the assay plate and allowed to lay eggs for 8 h (datasets 1 and 3) or 15 h (dataset 2) at 23.5 °C, then removed. Incubation of plates at 23.5 °C was continued until the 0 mM SPD N2 plates contained many mid-late stage L4s, at which point worms were rinsed off the plates and stored in microfuge tubes at -20 °C. Worm lengths were determined by mounting the thawed (dead) animals on an agarose pad covered by a coverslip, capturing images using a Leica M205 FA microscope equipped with digital camera plus software, and measuring the length of each worm from snout to tail tip using ImageJ.

C. elegans expresses three ATP13A2 orthologues: CATP-5, CATP-6 and CATP-7. Therefore, in this study the following mutant alleles were used, each of which was backcrossed to N2 Bristol at least three times: *catp-6(ok3473)* IV, *catp-7(tm4438)* IV, *catp-5(tm4481)* X⁵⁶. Each of these is a null allele, so they are referred to in the text as *catp-#(0)*.

Transgenic strains carrying extrachromosomal arrays with the pRF4 *rol-6(su1006)* plasmid were generated by microinjection and maintained as described previously⁵⁷. Injection mixes typically contained 100 μ g ml⁻¹ pRF4, plus 30 μ g ml⁻¹ of the test construct. In some cases, *Pmyo-2::gfp* was included at approximately 5 μ g ml⁻¹ as an additional method for detecting transgenic animals. To generate transport-defective versions of *catp-6* and *catp-7*, plasmids carrying *catp-6::mKate2* and *catp-7::GFP*²² were modified to change the coding sequence for the conserved DKTGT autophosphorylation motif to NKTGT. Proper expression and subcellular localization of CATP-6::mKate2 and CATP-7::GFP was verified in the case of all transgenic strains that were used. (*catp-6(0);[catp-6(+)]*) and *catp-7(0);[catp-7(+)]* express CATP-6 and wild-type CATP-7 in the null background, respectively; *catp-6(0);[catp-6(D465N)]* and *catp-7(0);[catp-7(D503N)]* express CATP-6(D465N) and CATP-7(D503N) in the null background).

Statistics and reproducibility

Data are expressed as the mean \pm s.e.m. or with individual data points (replicates of multiple independent experiments) shown on group means or box and whisker plots (indication of median, 25th percentile, 75th percentile and minimum to maximum value range). Flow cytometry was set up and gated as described in Supplementary Fig. 2. GraphPad Prism 7.04 was used to plot all graphs and to perform all of the required statistical and quantitative assessments. Statistical tests for each graph are described in the legend. The number of independent biological experiments for each panel is highlighted in the figure legends. For the cell biological experiments using the SH-SY5Y control, ATP13A2 KO, KO/WT and KO/D508N cell lines, each cell model is the sum of two independent clones, each performed a minimum of three independent times. For the quantification of immunoblots and radiograms, ImageJ and ImageQuant programmes were used. Experiments on various model systems were executed by different researchers, which provided consistent results that independently confirmed the major conclusions.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Gel source data for immunoblots and radiograms (Figs. 1, 2, Extended Data Figs. 1–3, 6) are available with the online version of the paper (Supplementary Fig. 1). All other datasets generated within this study are presented and analysed within this manuscript and are available from the corresponding author upon reasonable request.

31. Axelsen, K. B. & Palmgren, M. G. Evolution of substrate specificities in the P-type ATPase superfamily. *J. Mol. Evol.* **46**, 84–101 (1998).
32. Grunewald, A. et al. ATP13A2 mutations impair mitochondrial function in fibroblasts from patients with Kufor-Rakeb syndrome. *Neurobiol. Aging* **33**, 1843.e1–1843.e7 (2012).
33. Usenovic, M., Tresse, E., Mazzulli, J. R., Taylor, J. P. & Krainc, D. Deficiency of ATP13A2 leads to lysosomal dysfunction, α -synuclein accumulation, and neurotoxicity. *J. Neurosci.* **32**, 4240–4246 (2012).
34. Santoro, L. et al. Novel ATP13A2 (PARK9) homozygous mutation in a family with marked phenotype variability. *Neurogenetics* **12**, 33–39 (2011).
35. Funayama, M. et al. Rapid screening of ATP13A2 variant with high-resolution melting analysis. *Mov. Disord.* **25**, 2434–2437 (2010).
36. Lin, C. H. et al. Novel ATP13A2 variant associated with Parkinson disease in Taiwan and Singapore. *Neurology* **71**, 1727–1732 (2008).
37. Mao, X. Y. et al. ATP13A2 G2236A variant is rare in patients with early-onset Parkinson's disease and familial Parkinson's disease from mainland China. *Parkinsonism Relat. Disord.* **16**, 235–236 (2010).
38. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).
39. Kleinstiver, B. P. et al. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).
40. Doench, J. G. et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR–Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).
41. Covy, J. P., Waxman, E. A. & Giasson, B. I. Characterization of cellular protective effects of ATP13A2/PARK9 expression and alterations resulting from pathogenic mutants. *J. Neurosci. Res.* **90**, 2306–2316 (2012).
42. Gittler, A. D. et al. α -synuclein is part of a diverse and highly conserved interaction network that includes PARK9 and manganese toxicity. *Nat. Genet.* **41**, 308–315 (2009).
43. Kong, S. M. et al. Parkinson's disease-linked human PARK9/ATP13A2 maintains zinc homeostasis and promotes α -synuclein externalization via exosomes. *Hum. Mol. Genet.* **23**, 2816–2833 (2014).
44. Lambie, E. J., Tieu, P. J., Lebedeva, N., Church, D. L. & Conrad, S. CATP-6, a *C. elegans* ortholog of ATP13A2/PARK9, positively regulates GEM-1, an SLC16A transporter. *PLoS ONE* **8**, e77202 (2013).
45. Tsunemi, T. et al. Increased lysosomal exocytosis induced by lysosomal Ca^{2+} channel agonists protects human dopaminergic neurons from α -synuclein toxicity. *J. Neurosci.* **39**, 5760–5772 (2019).
46. Gietz, R. D. & Woods, R. A. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol.* **350**, 87–96 (2002).
47. Azouaoui, H. et al. A high-yield co-expression system for the purification of an intact Drs2p-Cdc50p lipid flippase complex, critically dependent on and stabilized by phosphatidylinositol-4-phosphate. *PLoS ONE* **9**, e112176 (2014).
48. Jidenko, M., Lenoir, G., Fuentes, J. M., le Maire, M. & Jaxel, C. Expression in yeast and purification of a membrane protein, SERCA1a, using a biotinylated acceptor domain. *Protein Expr. Purif.* **48**, 32–42 (2006).
49. Papadopoulos, A. et al. Flippase activity detected with unlabeled lipids by shape changes of giant unilamellar vesicles. *J. Biol. Chem.* **282**, 15559–15568 (2007).
50. Byun, J. A. et al. Analysis of polyamines as carbamoyl derivatives in urine and serum by liquid chromatography–tandem mass spectrometry. *Biomed. Chromatogr.* **22**, 73–80 (2008).
51. Stroikin, Y., Mild, H., Johansson, U., Roberg, K. & Ollinger, K. Lysosome-targeted stress reveals increased stability of lipofuscin-containing lysosomes. *Age* **30**, 31–42 (2008).
52. Nilsson, C., Kagedal, K., Johansson, U. & Ollinger, K. Analysis of cytosolic and lysosomal pH in apoptotic cells by flow cytometry. *Methods Cell. Sci.* **25**, 185–194 (2003).
53. Osório, L. et al. Viral vectors expressing a single microRNA-based short-hairpin RNA result in potent gene silencing in vitro and in vivo. *J. Biotechnol.* **169**, 71–81 (2014).
54. Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).
55. Casadaban, M. J., Martinez-Arias, A., Shapira, S. K. & Chou, J. β -galactosidase gene fusions for analyzing gene expression in *Escherichia coli* and yeast. *Methods Enzymol.* **100**, 293–308 (1983).
56. The *C. elegans* Deletion Mutant Consortium. Large-scale screening for targeted knockouts in the *Caenorhabditis elegans* genome. *G3* **2**, 1415–1425 (2012).
57. Mello, C. C., Kramer, J. M., Stinchcomb, D. & Ambros, V. Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970 (1991).
58. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725–738 (2010).
59. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40 (2008).
60. Omasits, U., Ahrens, C. H., Müller, S. & Wollscheid, B. Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* **30**, 884–886 (2014).

Acknowledgements This work was funded by the Fonds Wetenschappelijk Onderzoek (FWO, Research Foundation Flanders) (G094219N to P.V., G092714 and G080517N to V. Baekelandt, SBO Neuro-TRAFFIC S006617N to V. Baekelandt, P.V., W.A. and J.V.S. and 1503117N to S.M.), the KU Leuven (OT/13/091; LysoCaN C16/15/073 to P.V., J.V.S., W.A. and P.A.; OT/14/120 to V. Baekelandt) and the Queen Elisabeth Medical Foundation for Neurosciences (P.V. and V. Baekelandt). S.v.V. is an aspirant FWO research fellow (11Y7518N) and J.P.K. is the recipient of a strategic basic research doctoral grant of the FWO (1S18518N). J.L. received Lundbeck postdoctoral fellowships R209-2015-2704 and R171-2014-663. Funding for E.L. and J.Z. was provided by DFG grant LA3380/2-1. Some nematode strains were provided by the *Caenorhabditis* Genetics Center, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). We thank P. P. Van Veldhoven (KU Leuven) for discussions; P. Chaltin (Center for Drug Design and Discovery) for financial support for V. Benoy; and M. Schuermans, T. Arslan, A. Florizoone, J. Van Asselberghs, J. Chen and N. Hamouda for technical assistance. We also acknowledge our frequent use of the facilities and equipment of the Leuven Viral Vector Core facility (KU Leuven), Mass Spectrometry Core (R. Derua, KU Leuven), Cell and Tissue Imaging Cluster (P. Vanden Bergh, KU Leuven) and the FACS Core (KU Leuven/VIB).

Author contributions The study was designed by P.V., S.M., S.v.V. and J.E.; C.V.d.H. generated all stable cell lines and analysed ATP13A2 knockdown; J.-P.D. performed and analysed the CRISPR experiments; R.V. and J.P.K. synthesized BODIPY-labelled polyamine analogues; J.P.K. performed and analysed acidic nanoparticle preparations; S.v.V. performed and analysed the biochemical experiments (except Extended Data Fig. 2f by V. Benoy); J.L. contributed to the ATP13A2 purification protocol; S.M. performed and analysed the cell biology experiments; B.G. performed metabolomics; G.G. isolated mouse neurons; E.L. and J.Z. performed *C. elegans* experiments; and P.V., S.v.V. and S.M. wrote the manuscript, which was reviewed by all authors.

Competing interests Patent WO-2018002350-A1 of KU Leuven describes methods for detecting compounds with therapeutic use that target ATP13A2 or related isoforms using biological material and assays described in the current manuscript. A second patent of KU Leuven describing ATP13A2 cell models described in this manuscript has also been filed.

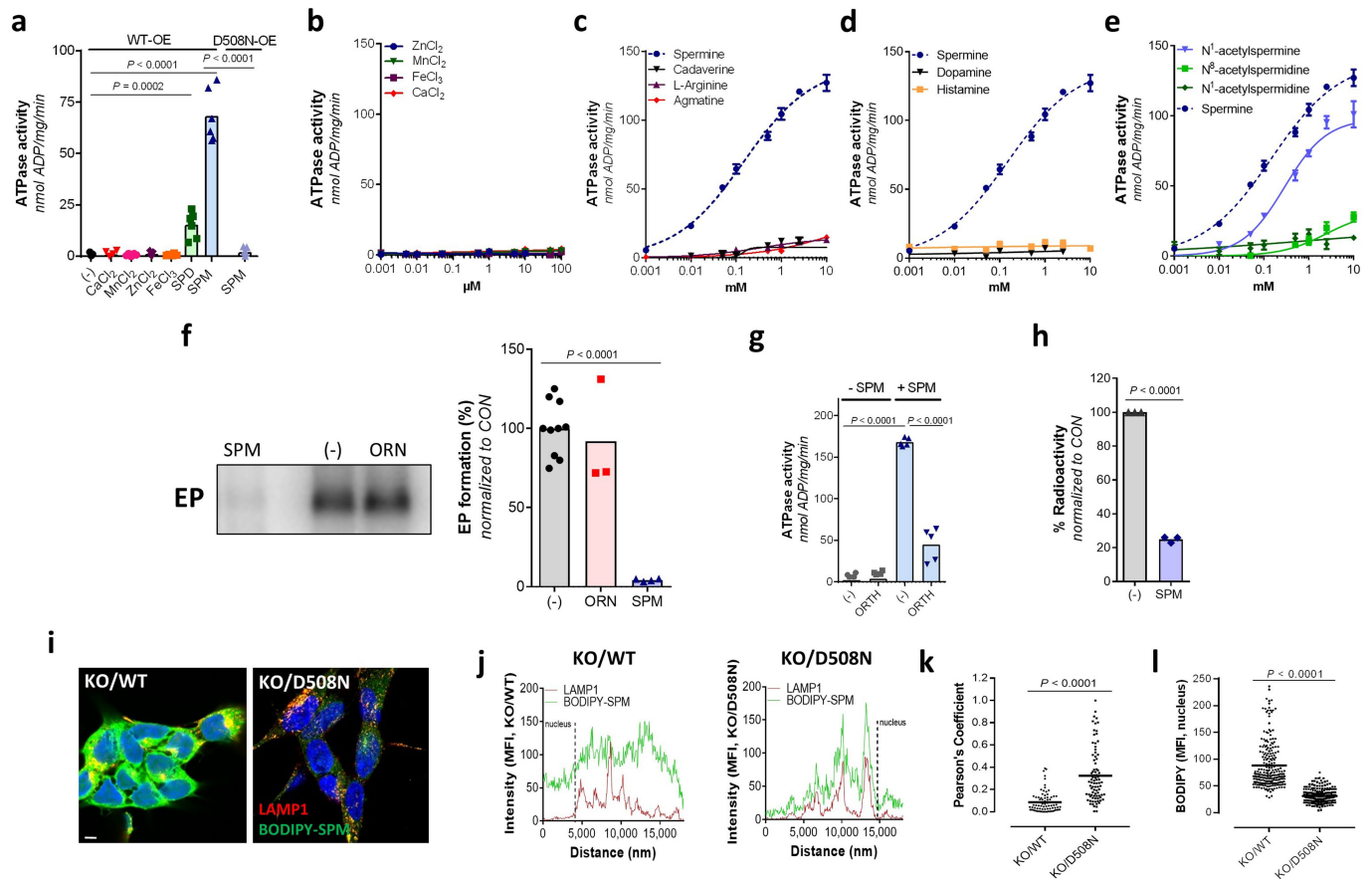
Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-1968-7>.

Correspondence and requests for materials should be addressed to P.V.

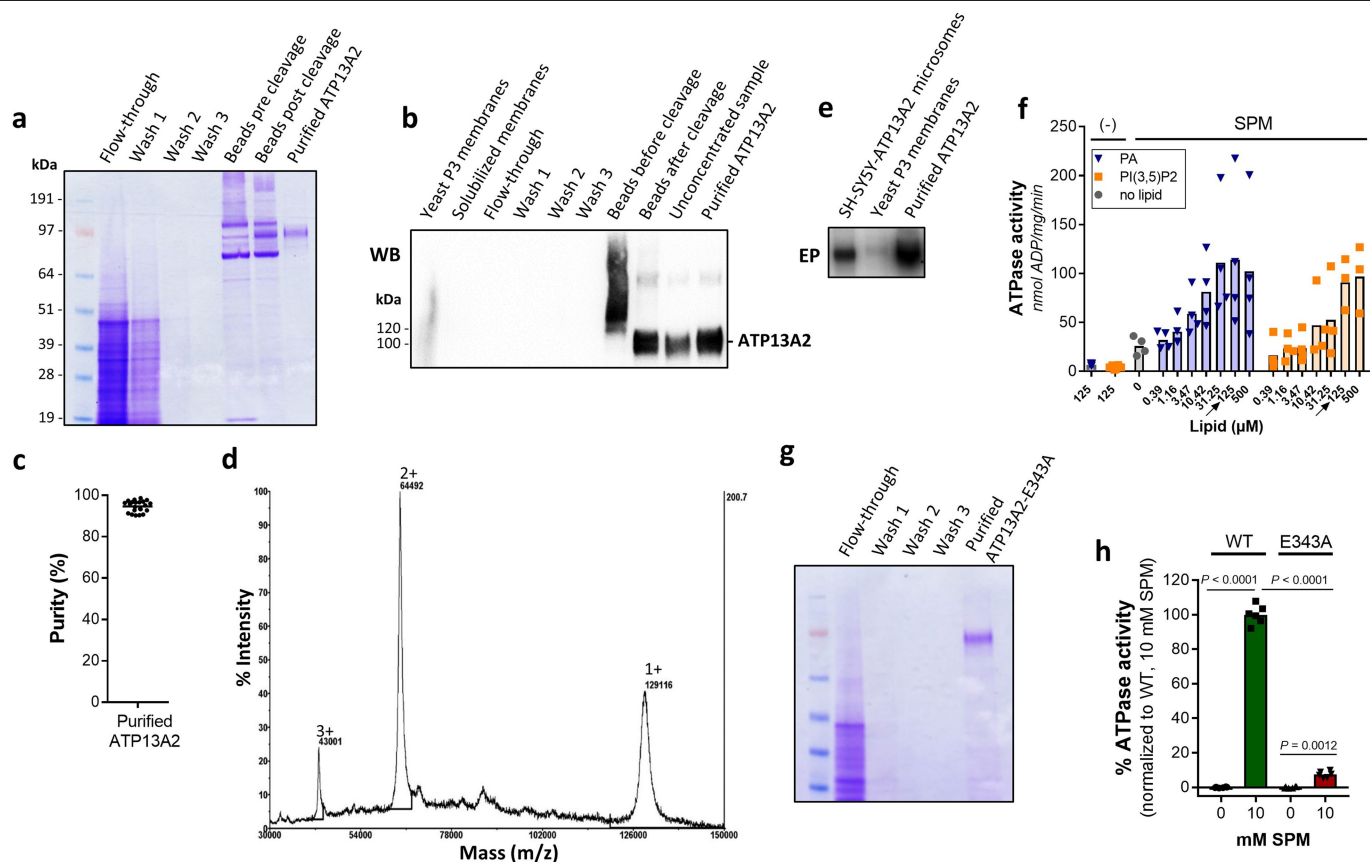
Peer review information Nature thanks Guillaume Lenoir, Hanne Poulsen, Taiji Tsunemi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | ATP13A2 is a polyamine transporter. **a–d**, The ATPase activity of ATP13A2 was measured in solubilized microsomes (5 μ g) collected from SH-SY5Y cells stably overexpressing wild-type ATP13A2 (WT-OE) in the presence of 100 nM CaCl₂, MnCl₂, ZnCl₂ or FeCl₃ and 100 μ M SPD or SPM (D508N-OE as a negative control, wild type was referenced from Fig. 1b) (**a**) or in the presence of the indicated doses of inorganic ions and heavy metals CaCl₂, MnCl₂, ZnCl₂ or FeCl₃ (**b**), diamines (cadaverine, agmatine and the amino acid L-arginine) (**c**), monoamines (dopamine and histamine) (**d**) and acetylated polyamines (N¹-acetylspermine, N⁸-acetylspermidine or N¹-acetylspermidine) (**e**). As a reference for **c–e**, we plotted the dose–response curve of SPM from Fig. 1b. **f**, Microsomes (20 μ g) collected from SH-SY5Y cells that overexpress ATP13A2 were incubated for 60 s with [γ -³²P]ATP in the presence of 10 mM ornithine (ORN) or SPM (referenced from Fig. 1c). Left, a representative autoradiogram of the phosphoenzymes (EP); right, quantification. CON, control. **g**, The ATPase activity of purified ATP13A2 was assessed after 1 mM SPM was administered in the presence or absence of 0.25 mM orthovanadate (ORTH), a general P-type ATPase inhibitor (supplemented with 125 μ M phosphatidic acid/PtdIns(3,5)P₂; conditions (–) and (–)/SPM refer to Fig. 1f).

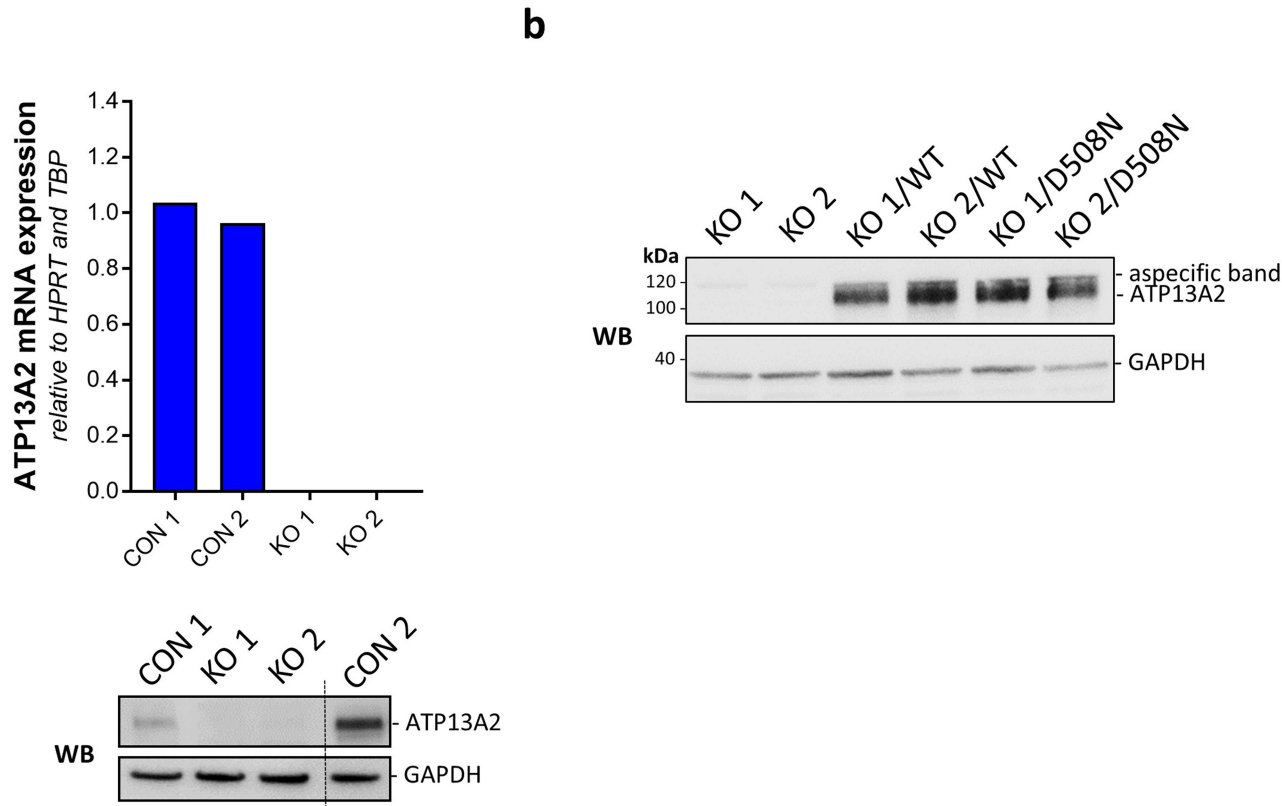
h, Purified ATP13A2 was incubated with [γ -³²P]ATP in the presence of 1 mM SPM, and radioactivity of the phospho-intermediate was assessed by scintillation counting. **i**, Comparison of the pulse (5 μ M, 15 min) chase (105 min, medium) BODIPY–SPM uptake in KO/WT and KO/D508N cell lines by confocal microscopy. Cells were subsequently stained with LAMP1 and imaged with the same laser settings by confocal microscopy. DAPI was used to visualize the nuclei. Scale bar, 5 μ m. **j**, Line intensity plots of the indicated dashed lines in Fig. 2d. **k**, Analysis of the Pearson's coefficient of LAMP1 and BODIPY–SPM for the images in Fig. 2d (KO/WT, 78 images; KO/D508N, 85 images). **l**, Mean fluorescence intensities (MFI) of BODIPY in DAPI-positive regions of samples shown in Fig. 2d (KO/WT, 233 nuclei; KO/D508N, 243 nuclei). Data are presented as the mean \pm s.e.m. or mean with individual data points shown (points represent replicates), with $n = 3$ independent biological experiments. Analysis was carried out using one-way ANOVA with Dunnett's (**f**) or Tukey's (**a, g**) corrections, or by two-tailed t -tests (unpaired, **h** or Welch's, **k, l**). Fitted lines are semi-log lines (**b**) or nonlinear allosteric sigmoidal association (**c–e**). For gel source data, see Supplementary Fig. 1.



Extended Data Fig. 2 | Streptavidin-based purification of wild-type ATP13A2

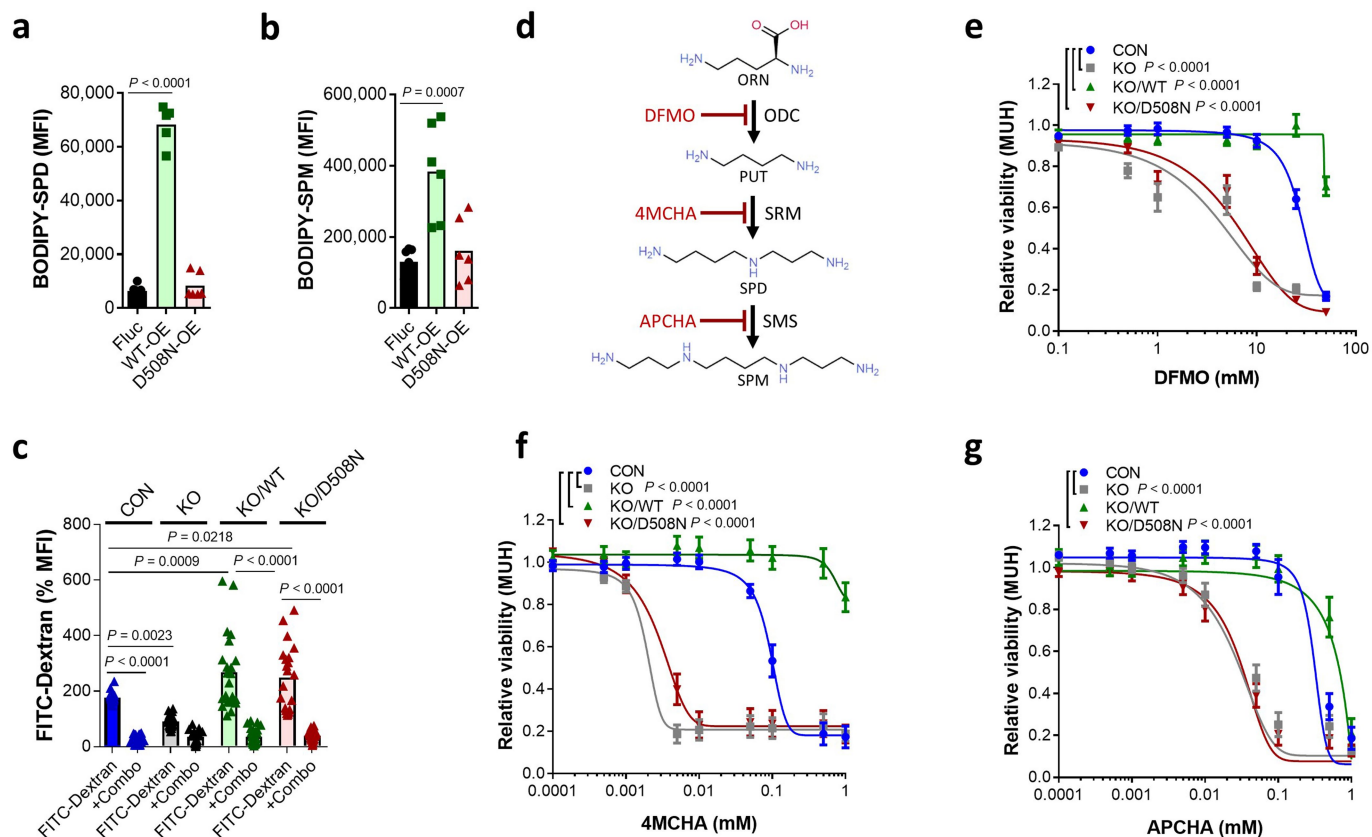
and the catalytically dead E343A mutant. **a**, Coomassie staining showing the purification process for wild-type ATP13A2, starting from solubilized yeast membrane fractions, followed by streptavidin affinity chromatography and on-column thrombin cleavage to elute the protein. **b**, Western blot analysis of stages in the purification of ATP13A2. **c**, Bar graph depicting protein purity as determined by densitometry from Coomassie-stained SDS-PAGE. **d**, Mass spectrometry analysis of the purified ATP13A2 sample. Singly, doubly and triply charged species are indicated. **e**, To evaluate phosphoenzyme formation, yeast P3 membranes (20 µg) and purified ATP13A2 (1 µg) were incubated for 60 s with [γ - 32 P]ATP. As a positive control, microsomes collected from SH-SY5Y cells that overexpress wild-type ATP13A2 (20 µg) were used. The image is a

representative radiogram depicting the ATP13A2 phosphoenzyme. **f**, The ATPase activity of purified ATP13A2 (0.3 µg) was measured in the presence of 2 mM SPM and the indicated concentrations of the ATP13A2 regulatory lipids phosphatidic acid (PA) and PtdIns(3,5)P₂. **g**, Coomassie staining showing the purification process for ATP13A2(E343A). **h**, The ATPase activity of purified wild-type ATP13A2 or ATP13A2(E343A) (0.5 µg) was measured in the presence of the indicated concentrations of SPM with 125 µM phosphatidic acid and 125 µM PtdIns(3,5)P₂. Data are expressed as mean with individual data points (points represent replicates) (**c**, **h**). The number of independent biological experiments were as follows: $n = 3$ (**b**, **e**–**h**); $n = 6$ (**d**); $n = 22$ (**a**, **c**). Analysis was performed using one-way ANOVA with Tukey's post-hoc correction (**h**). For gel source data, see Supplementary Fig. 1.



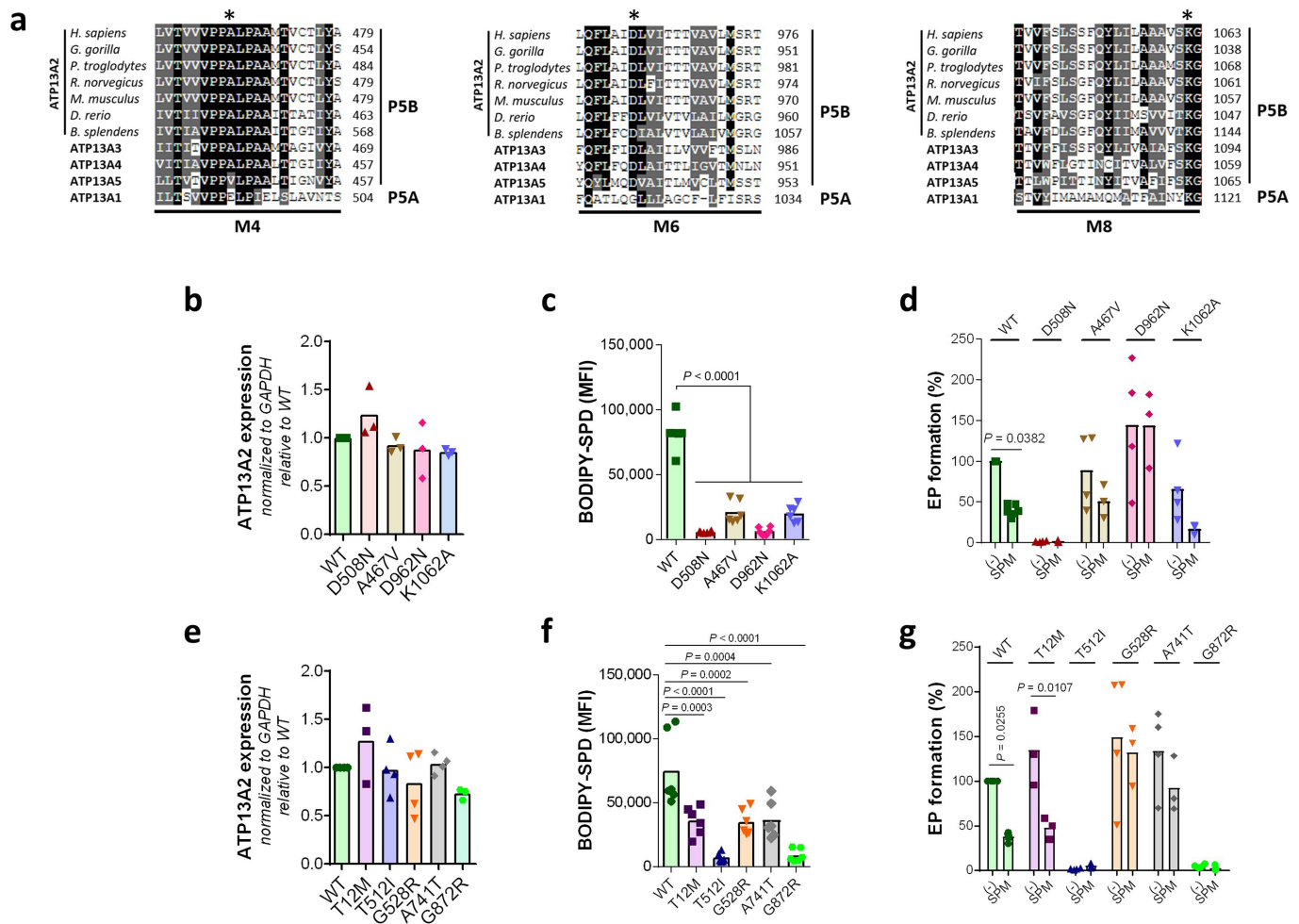
Extended Data Fig. 3 | Confirmation of CRISPR–Cas9-mediated ATP13A2 knockout and subsequent rescue with wild-type ATP13A2 or the D508N mutant. a, The ATP13A2 knockout cell lines (KO) were generated by CRISPR–Cas9 in SH-SY5Y cells and confirmed by qPCR (top) and immunoblotting (bottom). *Atp13a2* mRNA expression was normalized to hypoxanthine phosphoribosyltransferase (HPRT) and TATA-sequence-binding protein (TBP), and GAPDH was used as a loading control for the ATP13A2 protein levels. Two

fragments of the same blot are depicted and separated by a dotted line. **b**, Rescue of ATP13A2 knockout was performed by lentiviral transduction to express wild-type ATP13A2 (KO/WT) and the catalytically dead mutant D508N, which was used as a negative control (KO/D508N). The expression of the ATP13A2 constructs was confirmed via immunoblotting. The number of biologically independent experiments were as follows: $n = 1$ (**a**, top panel); $n = 3$ (**a**, bottom panel, **b**). For gel source data, see Supplementary Fig. 1.



Extended Data Fig. 4 | Polyamine uptake by ATP13A2 complements cytosolic polyamine synthesis. **a–c**, Assessment of the cellular uptake of BODIPY-labelled polyamine analogues (**a**, **b**) or FITC-dextran (**c**) by flow cytometry. Uptake of BODIPY-SPD (**a**) or BODIPY-SPM (**b**) in SH-SY5Y cells overexpressing Fluc (negative control), wild-type ATP13A2 (WT-OE) or the catalytically dead mutant D508N (D508N-OE). The cells were incubated with 5 μ M BODIPY-SPM or BODIPY-SPD for 2 h before analysis by flow cytometry. **c**, Analysis of FITC-dextran uptake (as a measure of endocytic capacity) was performed in SH-SY5Y control (CON) cells with endogenous ATP13A2 expression, ATP13A2 knockout cells (KO) and rescue cell lines with expression of wild-type ATP13A2 (KO/WT) or the D508N mutant (KO/D508N). The cells were pre-treated for 30 min with a combination (combo) of endocytosis inhibitors Dynasore (100 μ M), genistein (50 μ M) and Pitstop 2 (50 μ M). The cells were incubated for an additional 2 h

with FITC-dextran at 37 °C, followed by flow cytometry. **d**, Schematic representation of polyamine synthesis. SRM, spermidine synthase; SMS, spermine synthase. Specific inhibitors are indicated in red. Control, KO, KO/WT or KO/D508N cells were subjected to inhibition of polyamine synthesis by DFMO (**e**), 4MCHA (**f**) and APCHA (**g**) before measuring cell viability via the MUH assay. All data represent the average of two independent CRISPR-Cas9 knockout and control clones. All data are presented as mean with data points overlaid (points represent replicates) or mean \pm s.e.m. The number of biologically independent experiments were as follows: $n = 3$ (**a**, **b**); $n = 4$ (**c**, **e–g**). Analysis was performed using one-way ANOVA with Dunnett's (**a**, **b**) or Tukey's (**c**) post-hoc correction or two-way ANOVA with Dunnett's post-hoc correction (**e–g**). Fitted lines indicate nonlinear log(inhibitor) versus response (variable slope) (**e–g**).



Extended Data Fig. 5 | Catalytic and clinical mutations of ATP13A2 perturb polyamine function. **a**, Sequence alignment of predicted transmembrane helices M4 (left), M6 (middle) and M8 (right). The alignment was generated using Clustal Omega. We generated mutants in M4 (A467V), M6 (D962N) and M8 (K1062A). The A467V mutation converts the protein sequence PPALP of the predicted substrate-binding site in transmembrane segment M4 into the protein sequence PPVLP that is present in ATP13A5^{4,30}. Neighbouring membrane helices also contribute to substrate coordination in P-type ATPases, which often relies on conserved and charged residues, such as D962 in M6 and K1062 in M8 of ATP13A2⁴. **b**, Densitometry of the expression of catalytic mutants presented in Fig. 2e. **c**, Flow cytometric analysis of cellular BODIPY-SPD uptake in SH-SY5Y cells overexpressing wild-type ATP13A2, the D508N mutant or catalytic mutants. **d**, Quantification of ATP13A2 phosphorylation levels (EP) presented in Fig. 2g. **e**, Densitometry analysis of the expression of

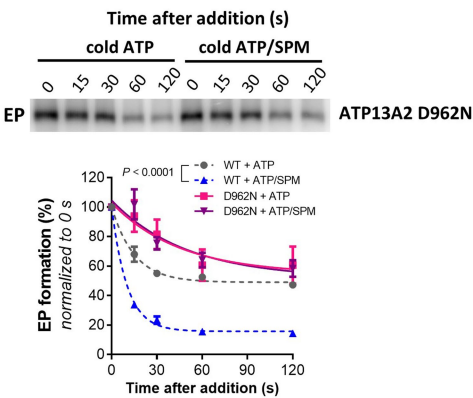
disease-related mutants presented in Fig. 2i. **f**, Flow cytometric analysis of cellular BODIPY-SPD uptake in SH-SY5Y cells overexpressing wild-type ATP13A2, D508N or disease mutants. **g**, Quantification of ATP13A2 phosphorylation levels presented in Fig. 2k. All data are depicted as mean with individual data points (points represent replicates). The number of independent biological experiments were as follows: **n** = 3 **b, c, d** (D508N (SPM), A467V (SPM), D962N (SPM), and K1062A (SPM)), **e** (T12M, G872R), **f, g** (T12M (-), T12M (SPM), T512I (SPM), G528R (SPM), A741T (SPM), and G872R (SPM)); **n** = 4 **d** (D508N (-), A467V (-), D962N (-), and K1062A (-)), **e** (wild-type, T512I, G528R and A741T), **g** (wild-type (-), T512I (-), G528R (-), A741T (-), and G872R (-)); **n** = 5 **d** (wild-type (SPM)) and **g** (wild-type (SPM)). Analysis by one-way ANOVA with Dunnett's (**b, c, e, f**) or two-way ANOVA with Sidak's (**d, g**) post-hoc corrections.

a

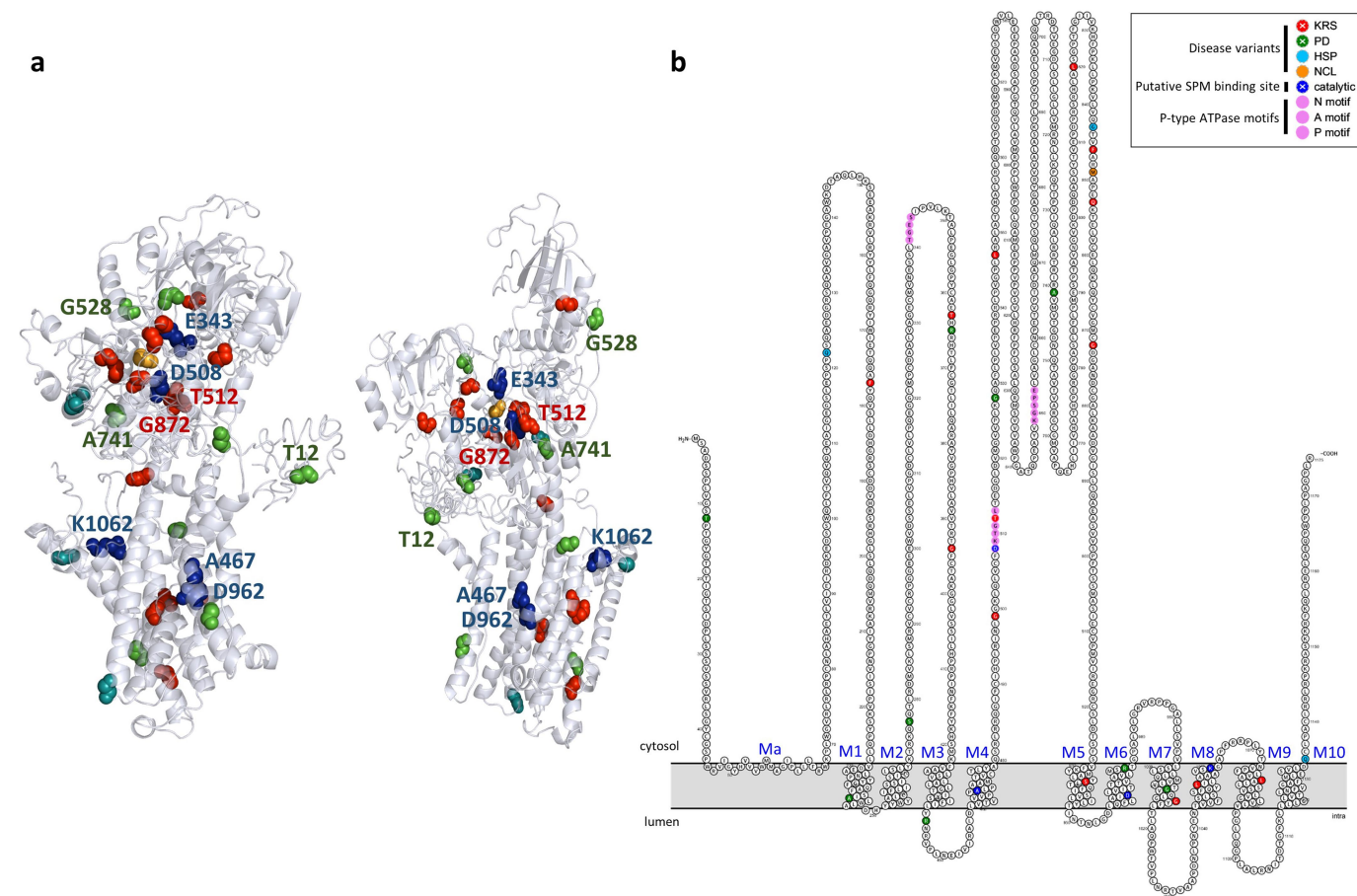
	Source	K ± SEM (s ⁻¹) ATP sensitivity	K ± SEM (s ⁻¹) ATP/SPM sensitivity
ATP13A2 WT	Microsomes	0.066 ± 0.010	0.098 ± 0.008
ATP13A2 D962N	Microsomes	0.022 ± 0.019	0.020 ± 0.012
ATP13A2 WT	Purified protein	0.050 ± 0.012	0.136 ± 0.030
ATP13A2 E343A	Purified protein	0.035 ± 0.031	0.039 ± 0.015

Extended Data Fig. 6 | The ATP sensitivity of ATP13A2(D962N) and ATP13A2(E343A) is independent of SPM. a, Overview of rate constants of ATP13A2 phosphoenzyme decay following a chase with non-radioactive ATP with or without 1 mM SPM. **b**, After 30 s of incubating D962N microsomes (20 µg) with [γ -³²P]ATP, the time course of dephosphorylation after an ATP chase was measured in the presence or absence of SPM. The top panel shows a

b

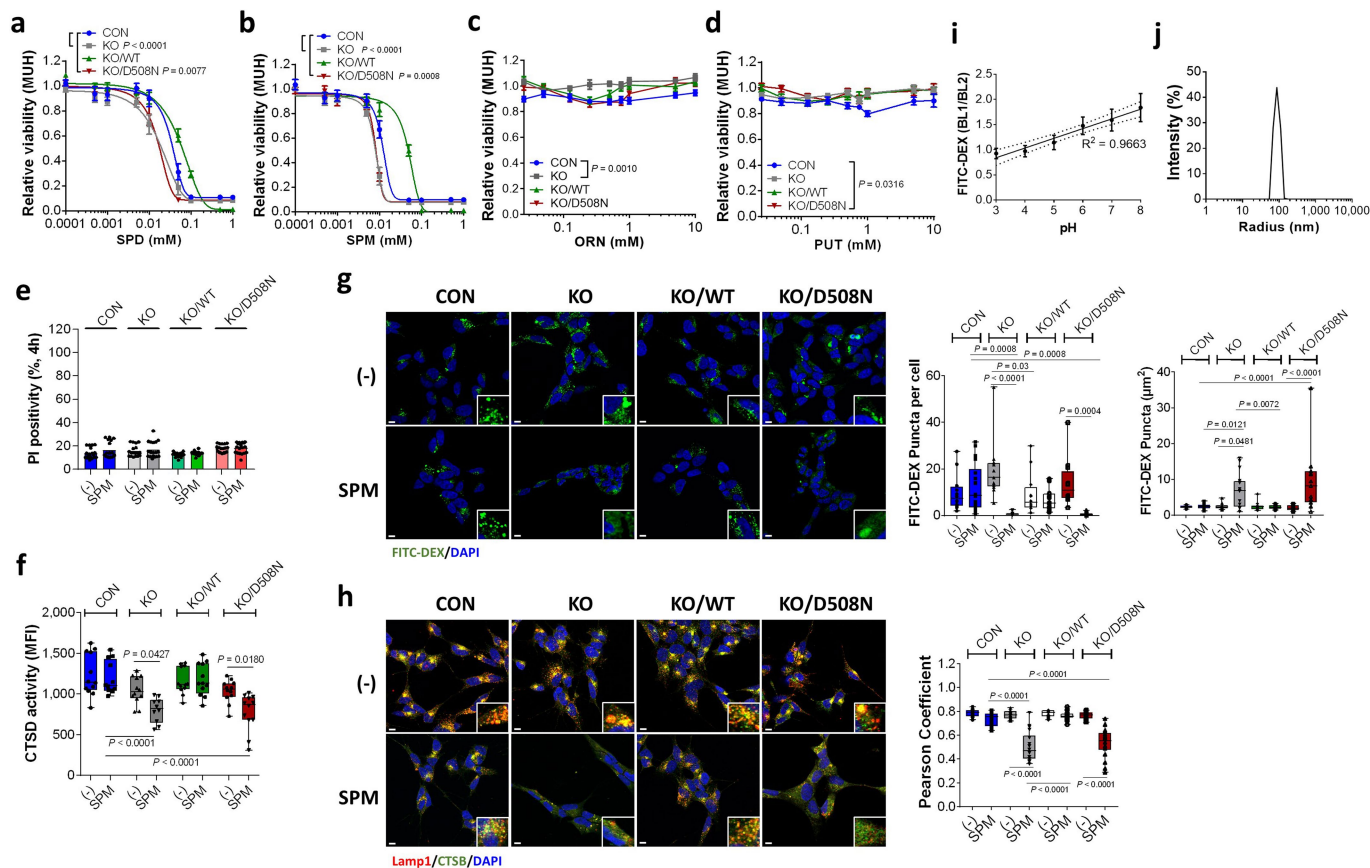


representative autoradiogram of the phosphoenzymes (EP), whereas the bottom panel depicts the quantification of ATP13A2 phosphorylation levels. As a reference we plotted the wild-type curve, shown in Fig. 1d. Data are presented as the mean ± s.e.m. of $n = 4$ biologically independent experiments. Analysis by two-way ANOVA with Tukey's test (**b**). The fitted line indicates two-phase decay (**b**). For gel source data, see Supplementary Fig. 1.



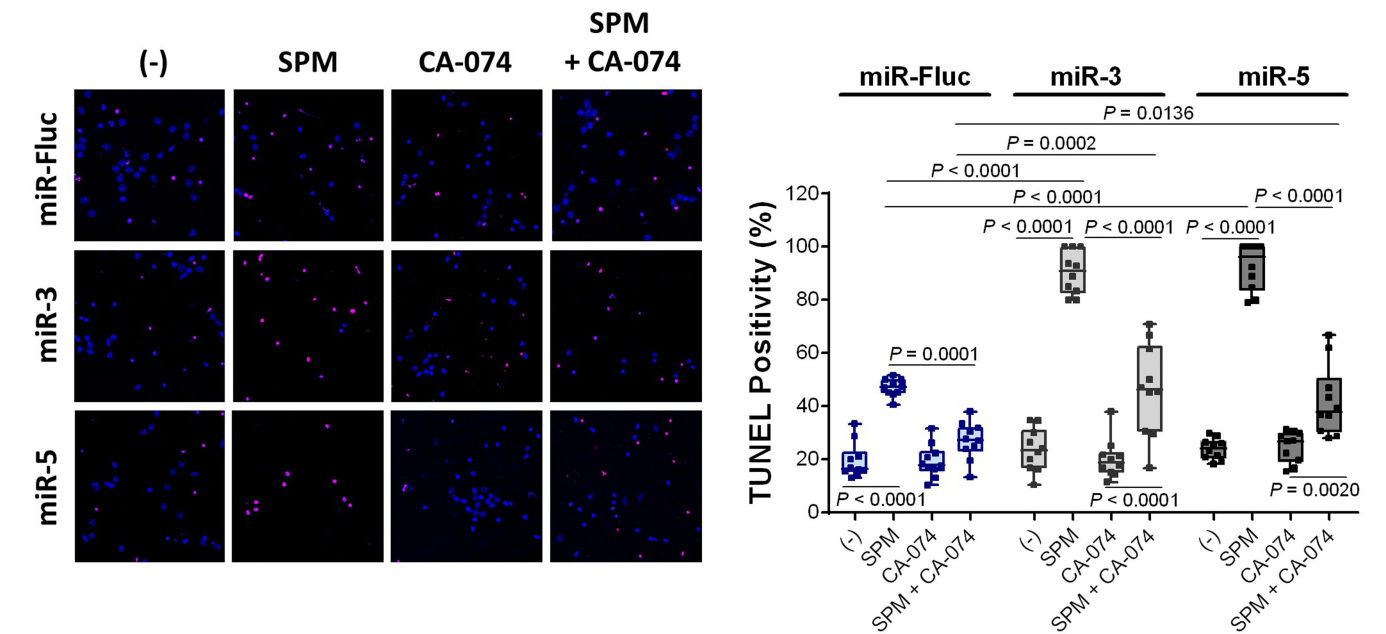
Extended Data Fig. 7 | Predicted topology of ATP13A2. a, Homology model of ATP13A2 based on the structure of Na⁺/K⁺-ATPase (ATP1A1, PDB ID: 3A3Y) as a template, generated by iTASSER (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>)^{58,59}. **b**, Predicted membrane topology of ATP13A2 visualized by Protter⁶⁰ (<http://wlab.ethz.ch/protter>). ATP13A2 consists of 10 transmembrane helices (M1–10) and an N-terminal membrane-associated region (Ma)⁹. Kufor-Rakeb syndrome (KRS)-associated mutations are highlighted in red; early-

onset Parkinson's disease (PD)-associated mutations in green; hereditary spastic paraplegia (HSP)-associated mutations in light blue; neuronal ceroid lipofuscinosis (NCL)-associated mutations in orange. Catalytic mutations and mutations in the predicted substrate-binding region are highlighted in dark blue. Residues that were subjected to mutagenesis in this study are labelled (only in **a**). P-type ATPase signature motifs in the cytosolic domains are indicated in pink (only in **b**).



Extended Data Fig. 8 | Lysosomal functionality and recovery. a–d. The impact of exogenous polyamines on cell toxicity (24 h) and lysosomal functionality (4 h) was assessed in SH-SY5Y control cells (CON) with endogenous ATP13A2, ATP13A2 knockout cells (KO) and rescue cell lines with wild-type expression (KO/WT) or expression of the catalytically dead mutant D508N (KO/D508N) on the KO background. Cytotoxicity of SPD (**a**), SPM (**b**), ORN (**c**) and PUT (**d**) were assessed via a MUH cell-viability assay. **e**, Death of the aforementioned cells was assessed after 4 h of SPM exposure (10 μ M) by propidium iodide (PI)-based flow cytometry. **f**, Measurement of cathepsin D activity. **g, h**, Lysosomal rupture under basal (–) and SPM (10 μ M) conditions was assessed via loss of FITC–dextran (FITC–DEX) punctae (**g**) or loss of cathepsin B (CTSB)/LAMP1 colocalization (**h**). Confocal images depict representative images with or without SPM exposure (4 h, DAPI staining for nuclei was included as a reference). Scale bars, 10 μ m. The box and whisker plots in **g** depict the frequency (left) and size (right) of FITC–DEX punctae; in **h**

the Pearson coefficient of colocalization of cathepsin B and LAMP1. **i**, Lysosomal pH (Fig. 3b) was evaluated using the fluorescent probe FITC–dextran and a dual-emission ratiometric technique. FITC is excited at 488 nm and emission is analysed at 530 nm (BL1) and 610 nm (BL2). A pH calibration curve was generated using FITC–dextran in cells permeabilized with 100 μ M monensin and equilibrated with calibration buffers (pH 3–8). **j**, Representative size distribution of the acidic nanoparticles used in this study. Data are presented as the mean \pm s.e.m. (**a–d, i**) or individual data points (representing replicates) overlaid on group means (**e**) or box and whisker plots (**f–h, line, median; box boundaries, 25th and 75th percentiles**). The number of independent biological experiments were as follows: $n = 3$ (**e–j**); $n = 4$ (**a–d**). Analysis was performed using two-way ANOVA with Dunnett’s (**a, b**) or Bonferroni’s (**c, d**) post-hoc corrections, or one-way ANOVA with Dunnett’s (**e**), Sidak’s (**f, g** (right)) or Tukey’s (**g** (left), **h**) post-hoc corrections. Fitted lines indicate nonlinear log(inhibitor) versus response (variable slope) (**a, b**).



Extended Data Fig. 9 | Inhibition of cathepsin B activity attenuates SPM-induced neuronal death. The effect of a cathepsin B inhibitor (CA-074, 10 μ M) on SPM-induced (10 μ M, 24 h) cell death in control (miR-Fluc) and *Atp13a2* knockdown (miR-3 and miR-5) neurons was assayed via TUNEL-based staining. Left, representative confocal images depicting TUNEL-positive neurons; right,

box and whisker plots with the quantification of the TUNEL staining. Data are presented as box and whisker plots (line, median; box boundaries, 25th and 75th percentiles) for which individual data points (representing replicates) are shown. $n = 3$ biologically independent experiments. Analysis by one-way ANOVA with Tukey's post-hoc correction.

Extended Data Table 1 | Apparent K_m and V_{max} values for ATP13A2 in the presence of various polyamines

	$K_m \pm \text{SEM} (\mu\text{M})$		$V_{max} \pm \text{SEM} (\text{nmol ADP/mg/min})$	
	Microsomal ATP13A2	Purified ATP13A2	Microsomal ATP13A2	Purified ATP13A2
Spermine	149 ± 34	76 ± 26	140 ± 6	159 ± 10
N ¹ -acetylspermine	286 ± 48	not tested	98 ± 5	not tested
Spermidine	~ 1700*	890 ± 501	~ 106*	194 ± 26

Data are derived from Fig. 1b, f and Extended Data Fig. 1e. Data represent $n = 3$ (spermidine, N¹-acetylspermine), or $n = 6$ (spermine) biologically independent experiments. *Estimated values (could not be accurately determined following fitting of the curve).

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- ☒ ☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- ☐ ☒ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☒ ☐ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Typhoon TM FLA 9500 (phosphorimaging)
Flexstation 3 microplate reader (fluorescent and luminescent assays)
Bio-Rad ChemiDoc (western blot imager)
QuantaSmart TM (v2.03) (scintillation counting)
Attune Cytometric Software (v2.1) (flow cytometry)
ZEN SP5 2012 (confocal imaging)

Data analysis

GraphPad Prism Version 7.04 was used for data representation and statistical analysis. Image J and Zen 2.3 Lite were used for image analysis. Image Quant TL (version 8.1) was used for analysis of gel-based auto-phosphorylation assays. Image Lab (5.2.1) was used to analyze Western blots. XCalibur Quan tool (version 4.2.28.14) was used for polyamine quantification (LC MS). Flowing Software (v2.5.1) was used for flow cytometry analysis. i-TASSER (Protein structure and function predictions, <https://zhanglab.ccmb.med.umich.edu/I-TASSER/>). PyMOL (structure visualization, version 2.3)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Source data for immunoblots and radiograms (Fig. 1-2, Extended data Fig. 1-3, 6) are available with the online version of the paper. All other datasets generated

within this study are presented and analysed within this manuscript and are available from the corresponding author upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to pre-determine sample size. All fundamental experiments were performed a minimum of three times (independent biological experiments), including all relevant controls, to allow for the generation of S.E.M. Confirmation of ATP13A2 knockout (SH-SY5Y) was performed only once by qRT-PCR and N = 3 by western blot analysis. Knockdown of ATP13A2 in neurons was only confirmed two times by qRT-PCR.
Data exclusions	No data were excluded.
Replication	All experimental findings were reproduced in several independent biological experiments (N) with multiple technical replicates. For each figure panel, the number of independent experiments N is indicated in the figure legends. Conclusions were independently confirmed in different model systems (in vitro, in cell lines, isolated neurons and in vivo) handled by multiple researchers and across several laboratories.
Randomization	Samples were not randomized, but appropriate controls were included in each figure. For in vivo studies (nematodes), animals were evenly distributed such that each group had a similar mean/density at the start of the study.
Blinding	Analyses were not blinded because experiments were performed and analyzed by the same researchers. Biochemistry, cell biology and in vivo research were performed independently by different researchers and their findings support one another providing independent confirmation.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used	<p>Primary antibodies used in this study with supplier, catalogue number and lot number, as well as the used dilution.</p> <p>SIGMA anti-ATP13A2 antibody; catalogue number A3361, lot # 076M4839, (Western Blotting, Rabbit, 1:1000) anti-GAPDH antibody; catalogue number G8795, lot # 076M4785V, (Western Blotting, Mouse, 1:5000)</p> <p>ABCAM anti-galectin-3 antibody; catalogue number ab2785, lot # GR3200865-11, (Immunofluorescence, Mouse, 1:200) anti-cathepsin B antibody; catalogue number ab58802, lot # GR3181508-13, (Immunofluorescence, Mouse, 1:100) anti-lamp1 antibody; catalogue number ab24170, lot # GR3235359-1, (Immunofluorescence, Rabbit, 1:200)</p>
Validation	<p>Concerning antibody specificity, we kindly refer to the supplier's websites and datasheets to find statements on specificity and citations for the use of the antibodies:</p> <p>anti-ATP13A2 antibody; https://www.sigmaaldrich.com/catalog/product/sigma/a3361?lang=en&region=BE anti-GAPDH antibody; https://www.sigmaaldrich.com/catalog/product/sigma/g8795?lang=en&region=BE anti-galectin-3 antibody; https://www.abcam.com/galectin-3-antibody-a3a12-ab2785.html</p>

anti-cathepsin B antibody; <https://www.abcam.com/cathepsin-b-antibody-ca10-ab58802.html>
 anti-lamp1 antibody; <https://www.abcam.com/lamp1-antibody-lysosome-marker-ab24170.html>

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	SH-SY5Y from ATCC (ATCC number CRL-2266™) (lot number 62431864) SH-SY5Y in house collection HEK-293T from ATCC (https://www.lgcstandards-atcc.org/Products/All/CRL-11268.aspx?geo_country=be)
Authentication	SH-SY5Y (ATCC); were certified by ATCC by STR genotype analysis (https://www.lgcstandards-atcc.org/en/Products/All/CRL-2266.aspx#documentation). SH-SY5Y cells; (from the in house collection) were authenticated via DNA fingerprinting (Leibniz-Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH). HEK-293T (ATCC); were certified by ATCC by STR genotype analysis (https://www.lgcstandards-atcc.org/Products/All/CRL-11268.aspx?geo_country=be#documentation)
Mycoplasma contamination	All cell lines were routinely tested for mycoplasma contamination using the MycoAlert Mycoplasma Detection kit (LT07-418) and no contaminations have been detected.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	Mice - E16 FVB/N mice (females for embryos, neuronal isolation). C. elegans - N2 Bristols (Males and females, larvae to adult growth analysis)
Wild animals	No wild animals were used within this research
Field-collected samples	No field collected samples were used within this research
Ethics oversight	All mouse experiments were carried out in accordance with the European Communities Council Directive of November 24, 1986 (86/609/EEC) and approved by the Bioethical Committee of the KU Leuven (Belgium) (ECD project P185-2014).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

- ☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☐ All plots are contour plots with outliers or pseudocolor plots.
- ☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Samples were harvested by TRYPLE, washed and stained. Prior to FACS, cells were resuspended in PBS containing BSA, filtered and stored on ice. Analysis was performed on living (non-fixed) cells.
Instrument	Attune NxT - Thermo Fisher Scientific
Software	Attune Cytometric Software (v2.1)
Cell population abundance	Flow cytometry was performed on cell lines and at no point tracked a subpopulation of a heterogeneous cell mixture. We assessed 10,000 events of the total cell population, including both living and dead cells.
Gating strategy	Total cells were gated (R1) to remove any signal contamination from debris. Upon selection live/dead cells were characterized in (RX/R1) or fluorescence acquired (R2). In all cases samples were gated on the position of the negative population (unstained) of each cell line. Detection was obtained within the relevant detection window.

- ☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Clades of huge phages from across Earth's ecosystems


<https://doi.org/10.1038/s41586-020-2007-4>

Received: 22 March 2019

Accepted: 2 January 2020

Published online: 12 February 2020

Open access

 Check for updates

Basem Al-Shayeb^{1,17}, Rohan Sachdeva^{1,17}, Lin-Xing Chen¹, Fred Ward¹, Patrick Munk², Audra Devoto¹, Cindy J. Castelle¹, Matthew R. Olm¹, Keith Bouma-Gregson³, Yuki Amano⁴, Christine He¹, Raphaël Méheust¹, Brandon Brooks¹, Alex Thomas¹, Adi Lavy¹, Paula Matheus-Carnevali¹, Christine Sun⁵, Daniela S. A. Goltsman⁵, Mikayla A. Borton⁶, Allison Sharrar³, Alexander L. Jaffe¹, Tara C. Nelson⁷, Rose Kantor¹, Ray Keren¹, Katherine R. Lane¹, Ibrahim F. Farag¹, Shufei Lei³, Kari Finstad⁸, Ronald Amundson⁸, Karthik Anantharaman³, Jinglie Zhou⁹, Alexander J. Probst¹, Mary E. Power¹⁰, Susannah G. Tringe⁹, Wen-Jun Li¹¹, Kelly Wrighton⁶, Sue Harrison¹², Michael Morowitz¹³, David A. Relman⁵, Jennifer A. Doudna¹, Anne-Catherine Lehours¹⁴, Lesley Warren⁷, Jamie H. D. Cate¹, Joanne M. Santini¹⁵ & Jillian F. Banfield^{1,3,8,16}✉

Bacteriophages typically have small genomes¹ and depend on their bacterial hosts for replication². Here we sequenced DNA from diverse ecosystems and found hundreds of phage genomes with lengths of more than 200 kilobases (kb), including a genome of 735 kb, which is—to our knowledge—the largest phage genome to be described to date. Thirty-five genomes were manually curated to completion (circular and no gaps). Expanded genetic repertoires include diverse and previously undescribed CRISPR–Cas systems, transfer RNAs (tRNAs), tRNA synthetases, tRNA-modification enzymes, translation-initiation and elongation factors, and ribosomal proteins. The CRISPR–Cas systems of phages have the capacity to silence host transcription factors and translational genes, potentially as part of a larger interaction network that intercepts translation to redirect biosynthesis to phage-encoded functions. In addition, some phages may repurpose bacterial CRISPR–Cas systems to eliminate competing phages. We phylogenetically define the major clades of huge phages from human and other animal microbiomes, as well as from oceans, lakes, sediments, soils and the built environment. We conclude that the large gene inventories of huge phages reflect a conserved biological strategy, and that the phages are distributed across a broad bacterial host range and across Earth's ecosystems.

Phages—viruses that infect bacteria—are considered distinct from cellular life owing to their inability to carry out most biological processes required for reproduction. They are agents of ecosystem change because they prey on specific bacterial populations, mediate lateral gene transfer, alter host metabolism and redistribute bacterially derived compounds through cell lysis^{2–4}. They spread antibiotic resistance⁵ and disperse pathogenicity factors that cause disease in humans and animals^{6,7}. Most knowledge about phages is based on laboratory-studied examples, the vast majority of which have genomes that are a few tens of kb in length. Widely used isolation-based methods select against large phage particles, and they can be excluded from phage concentrates obtained by passage through 100-nm or 200-nm filters¹. In 2017, only 93 isolated phages with genomes that were more than

200 kb in length were published¹. Sequencing of whole-community DNA can uncover phage-derived fragments; however, large genomes can still escape detection owing to fragmentation⁸. A new clade of human- and animal-associated megaphages was recently described on the basis of genomes that were manually curated to completion from metagenomic datasets⁹. This finding prompted us to carry out a more-comprehensive analysis of microbial communities to evaluate the prevalence, diversity and ecosystem distribution of phages with large genomes. Previously, phages with genomes of more than 200 kb have been referred to as ‘jumbophages’¹ or, in the case of phages with genomes of more than 500 kb, as megaphages⁹. As the set reconstructed here span both size ranges we refer to them simply as ‘huge phages’. A graphical abstract provides an overview of our approach and main

¹Innovative Genomics Institute, University of California Berkeley, Berkeley, CA, USA. ²National Food Institute, Technical University of Denmark, Kongens Lyngby, Denmark. ³Earth and Planetary Science, University of California Berkeley, Berkeley, CA, USA. ⁴Nuclear Fuel Cycle Engineering Laboratories, Japan Atomic Energy Agency, Tokai-mura, Japan. ⁵Department of Microbiology & Immunology, Stanford University, Stanford, CA, USA. ⁶Department of Soil and Crop Sciences, Colorado State University, Fort Collins, CO, USA. ⁷Department of Civil and Mineral Engineering, University of Toronto, Toronto, Ontario, Canada. ⁸Environmental Science, Policy and Management, University of California Berkeley, Berkeley, CA, USA. ⁹DOE Joint Genome Institute, Berkeley, CA, USA. ¹⁰Integrative Biology, University of California Berkeley, Berkeley, CA, USA. ¹¹School of Life Sciences, Sun Yat-Sen University, Guangzhou, China. ¹²Centre for Bioprocess Engineering Research, University of Cape Town, Cape Town, South Africa. ¹³Department of Surgery, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. ¹⁴Laboratoire Microorganismes: Génome et Environnement, Université Clermont Auvergne, CNRS, Clermont-Ferrand, France. ¹⁵Institute of Structural and Molecular Biology, University College London, London, UK. ¹⁶School of Earth Sciences, University of Melbourne, Melbourne, Victoria, Australia. ¹⁷These authors contributed equally: Basem Al-Shayeb, Rohan Sachdeva. ✉e-mail: jbanfield@berkeley.edu

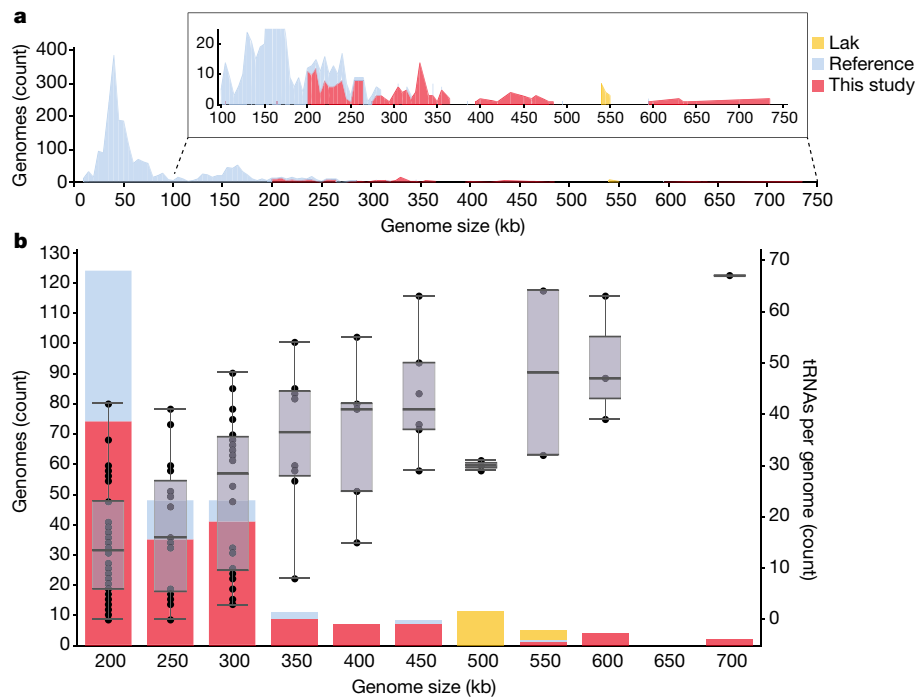


Fig. 1 | Distribution of the genome sizes and tRNAs of phages. **a**, Size distribution of circularized bacteriophage genomes from this study, Lak megaphage genomes reported recently for a subset of the same samples⁹ and reference sources. Reference genomes were collected from all complete RefSeq r92 dsDNA genomes and non-artefactual assemblies with lengths of more than 200 kb from a previous study¹⁴. **b**, Histogram of the genome size

distribution of phages with genomes of more than 200 kb from this study, Lak and reference genomes. Box-and-whisker plot of tRNA counts per genome from this study and Lak phages as a function of genome size (Spearman's $\rho = 0.61$, $P = 4.5 \times 10^{-22}$, $n = 201$ individual phage genomes). The middle line for each box marks the median tRNA count for each size bin, the box marks the interquartile range, and the whiskers represent the maximum and minimum.

findings (Extended Data Fig. 1). This study expands our understanding of phage biodiversity and reveals the wide variety of ecosystems in which phages have genomes with sizes that rival those of small-celled bacteria^{10–12}. We postulate that these phages have evolved a distinct ‘life’ strategy that involves extensive interception and augmentation of host biology while they replicate their huge genomes.

Ecosystem sampling

Metagenomic datasets were acquired from human faecal and oral samples, faecal samples from other animals, freshwater lakes and rivers, marine ecosystems, sediments, hot springs, soils, deep subsurface habitats and the built environment (Extended Data Fig. 2). Genome sequences that were clearly not bacterial, archaeal, archaeal virus, eukaryotic or eukaryotic virus were classified as phage, plasmid-like or mobile genetic elements of uncertain nature on the basis of their gene inventories (Supplementary Information). De novo assembled fragments close to or more than 200 kb in length were tested for circularization and a subset was selected for manual verification and curation to completion (Methods).

Genome sizes and basic features

We reconstructed 351 phage sequences, 6 plasmid-like sequences and 4 sequences of unknown classification (Extended Data Fig. 2). We excluded additional sequences that were inferred to be plasmids (Methods), retaining only those that encoded CRISPR–Cas loci. We included 3 phage sequences of ≤ 200 kb in length owing to the presence of CRISPR–Cas loci. Consistent with the classification as phages, we identified a wide variety of phage-relevant genes, including those involved in lysis and encoding structural proteins, and documented other expected

genomic features of phages (Supplementary Information). Some predicted proteins were large, up to 7,694 amino acids in length; some were tentatively annotated as structural proteins. In total, 175 phage sequences were circularized and 35 were manually curated to completion, in some cases by resolving complex repeat regions, revealing their encoded proteins (Methods and Supplementary Table 1). The remaining genomes are probably incomplete, although some may be complete, but linear. Approximately 30% of genomes show clear GC skew indicative of bidirectional replication and 30% have patterns indicative of unidirectional replication¹³ (Extended Data Fig. 3 and Supplementary Information).

Our 4 largest complete, manually curated and circularized phage genomes are 634, 636, 642 and 735 kb in length and are—to our knowledge—the largest phage genomes reported to date. The largest previously reported circularized phage genome was 596 kb in length¹⁴. The same previous study also reported a circularized genome of 630 kb in length; however, this is an assembly artefact (Supplementary Information). The problem of concatenation artefacts was sufficiently prominent in IMG/VR¹⁵ that we did not include these data in further analyses. We used both complete and circularized genomes from our study and published phage genomes to produce an updated view of the distribution of phage genome sizes (Methods). Without the huge phages reported here, the median genome size for complete phages is around 52 kb (Fig. 1a). Thus, the sequences reported here substantially expand the inventory of phages with unusually large genomes (Fig. 1b).

Some of our reported genomes have a very low coding density (9 genomes have densities of less than 78%) (Supplementary Information), probably owing to the use of a genetic code that is different from the standard code (Methods). This phenomenon has been rarely noted in phages, but has previously been reported for the Lak phage⁹ and in a previous study¹⁶. In the current study, some genomes (mostly those that are associated with humans and/or animals) appear to have reassigned

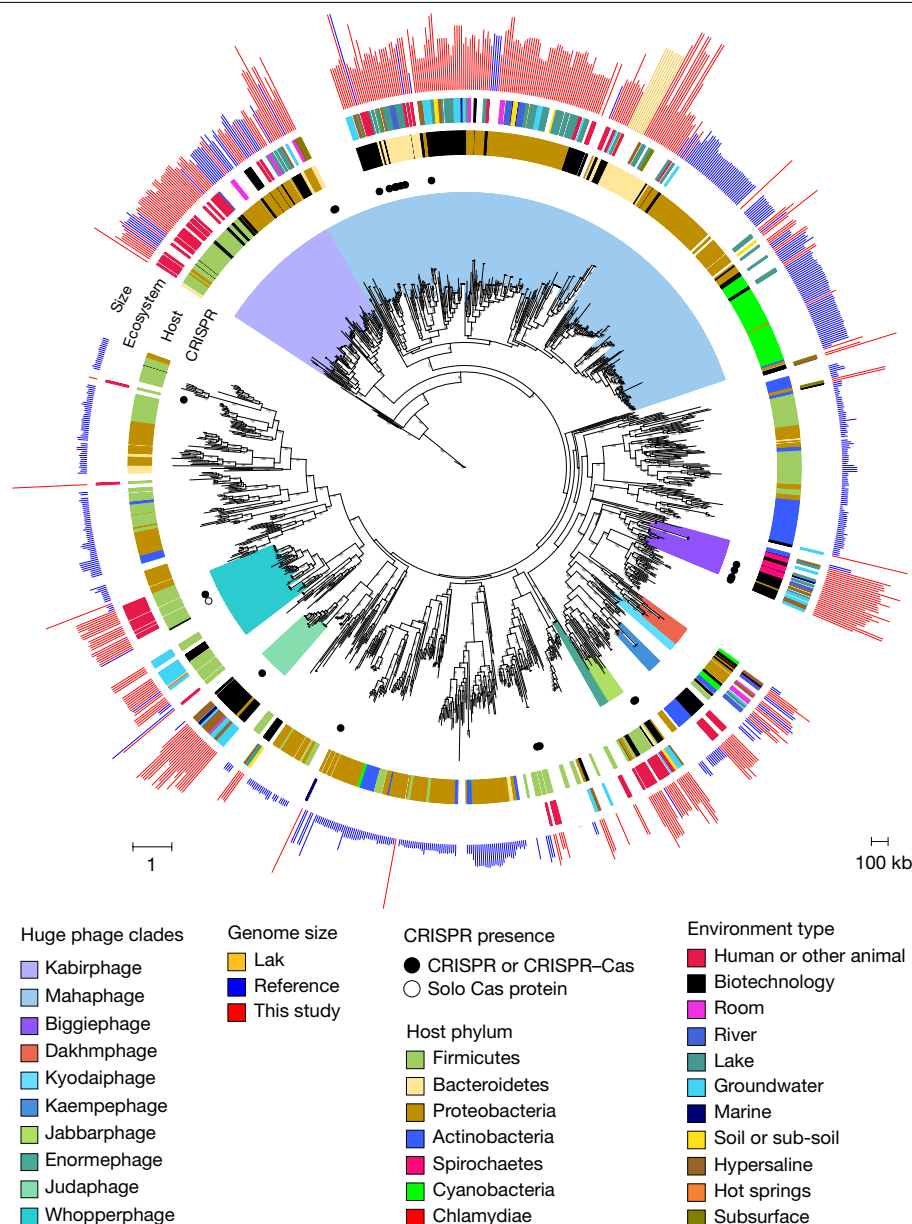


Fig. 2 | Phylogenetic reconstruction of the evolutionary history of huge phages. The phylogeny of phages was reconstructed using large terminase sequences from this study ($n = 397$) and similar matches from all RefSeq r92 proteins ($n = 532$). The tree also includes large terminase sequences from complete RefSeq phage, the Lak megaphage clade⁹ ($n = 9$) and non-artefactual phage genomes that are more than 200 kb, from a previous study¹⁴. Huge phage clades identified in this study were independently corroborated with a phylogenetic reconstruction of major capsid protein (MCP) genes (Extended

Data Fig. 5a) and protein clustering (Extended Data Fig. 5b). The tree was rooted using eukaryotic herpesvirus terminases ($n = 7$). The inner to outer rings display the presence of CRISPR-Cas in this study, host phylum, environmental sampling type and genome size. Host phylum and genome size were not included for RefSeq protein database matches for which the sequence may be from an integrated prophage or part of organismal genome projects. Scale bars show the number of substitutions per site (left) and number of base pairs (right).

the UAG (amber) stop codon to encode an amino acid (Extended Data Fig. 4 and Supplementary Information).

In only one case, we identified a sequence of more than 200 kb that was classified as a prophage on the basis of the transition into a flanking bacterial genome sequence. However, around half of the genomes were not circularized, so their potential integration as a prophage cannot be ruled out. The presence of integrases in some genomes is suggestive of a temperate lifestyle under some conditions.

Hosts, diversity and distribution

An intriguing question relates to the evolutionary history of phages with huge genomes; namely, whether they are the result of recent genome

expansion within clades of normal-sized phages or whether a large inventory of genes is an established, persistent strategy. To investigate this, we constructed phylogenetic trees for large terminase subunit proteins (Fig. 2) and major capsid proteins (Extended Data Fig. 5a) using sequences from public databases as a context (Methods). Many of the sequences from our phage genomes cluster together with high bootstrap support, thus defining clades. Analysis of the genome size information for database sequences shows that the public sequences that fall into these clades are from phages with genomes of at least 120 kb in length. The largest clade, referred to here as Mahaphage (Maha being Sanskrit for huge), includes all of our largest genomes as well as the 540–552 kb Lak genomes from human and animal microbiomes⁹. We identified nine other clusters of large phages, and refer to them

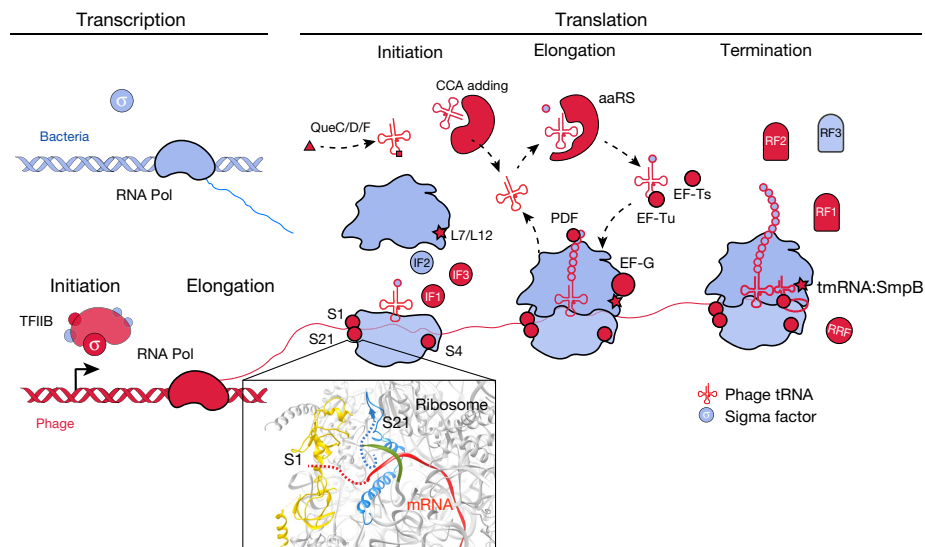


Fig. 3 | A model for phage interception and redirection of host translational systems. Potential mechanisms for how phage-encoded capacities could function to redirect the translational system of the host to produce phage proteins (bacterial components in blue, phage proteins in red). No huge phage encodes all translation-related genes, but many have tRNAs and tRNA synthetases (Supplementary Table 6). Phage proteins with up to six ribosomal protein S1 domains occur in a few genomes. The S1 binds to mRNA to bring it into the site on the ribosome where it is decoded³⁹. Phage ribosomal protein S21 might promote translation initiation of phage mRNAs, and many sequences

have N-terminal extensions that may be involved in binding RNA (dashed blue line in ribosome insert (RCSB Protein Data Bank (PDB) code: 6BU8⁴⁰)), analysed with UCSF Chimera⁴¹. Many other proteins of the translational apparatus that belong to all steps of the translation cycle are encoded by huge phages. aaRS, aminoacyl-tRNA synthetase; CCA-adding, tRNA nucleotidyltransferase; EF, elongation factor; IF, initiation factor; PDF, peptide deformylase; QueC/D/F, queuosine synthesis and tRNA modification; RF, release factor; RNA Pol, RNA polymerase; RRF, ribosome recycling factor; TFIIB, transcription factor IIB.

using the words for ‘huge’ in the languages of some authors of this paper. We acknowledge that the detailed tree topologies for different genes and datasets vary slightly; however, the clustering is broadly supported by protein family and capsid analyses (Extended Data Fig. 5a, b). The fact that large phages are consistently grouped together into clades establishes that a large genome size is a relatively stable trait. Within each clade, phages were sampled from a wide variety of environment types (Fig. 2), indicating the diversification of these huge phages and their hosts across ecosystems. We also examined the environmental distribution of phages that are so closely related that their genomes can be aligned and we found 20 cases in which the phages occur in at least 2 distinct cohorts or habitat types (Supplementary Table 2).

To determine the extent to which bacterial host phylogeny correlates with phage clades, we identified some phage hosts using CRISPR spacer targeting from bacteria in the same or related samples and phylogenies of normally host-associated phage genes (see below, Supplementary Table 3). We also tested the predictive value of bacterial taxonomic affiliations of the phage gene inventories (Methods) and found that in every case, CRISPR spacer targeting and phylogeny agreed with phylum-level taxonomic profiles. We therefore used taxonomic profiles to predict the bacterial host phylum for many phages (Supplementary Table 4). The results establish the importance of Firmicutes and Proteobacteria as hosts (Extended Data Fig. 2) ($P = 2.5 \times 10^{-5}$, $n = 74$, $W = 606$; one-sided Wilcoxon signed-rank test). The higher prevalence of Firmicutes-infecting huge phages in the human and animal gut compared with other environments reflects the potential host compositions of the microbiomes ($P = 9.3 \times 10^{-7}$, $n = 37$, $U = 238$; one-sided Mann–Whitney U -test). Notably, the 5 genomes that were more than 634 kb in length were all from phages that were predicted to replicate in Bacteroidetes, as do Lak phages⁹, and all cluster within the Mahaphage clade. Overall, phages that grouped together phylogenetically are predicted to replicate in bacteria of the same phylum (Fig. 2).

Metabolism, transcription and translation

The phage genomes encode proteins that are predicted to localize to the bacterial membrane or cell surface. These may affect the susceptibility of the host to infection by other phages (Supplementary Table 5 and Supplementary Information). We identified almost all of the previously reported categories of genes that have been suggested to augment host metabolism (Supplementary Information). Many phages have genes involved in the de novo biosynthesis of purines and pyrimidines, and the interconversion of nucleic and ribonucleic acids and nucleotide phosphorylation states. These gene sets are intriguingly similar to those of bacteria with very small cells and putative symbiotic lifestyles¹⁰ (Supplementary Table 5).

Notably, many phages have genes with predicted functions in transcription and translation (Supplementary Table 6). Complete phage genomes encode up to 67 tRNAs, with sequences that are distinct from those of their hosts (Supplementary Table 7). Generally, the number of tRNAs per genome increases with genome length (Fig. 1) (Spearman's $\rho = 0.61$, $P = 4.5 \times 10^{-22}$, $n = 201$). Huge phages have up to 15 tRNA synthetases per genome (Supplementary Table 7), which are also distinct from but related to those of their hosts (Extended Data Fig. 7a and Supplementary Information). Phages may use these proteins to charge their own tRNA variants with host-derived amino acids. A subset of genomes has genes for tRNA modification and ligation of tRNAs cleaved by host defenses.

Many phages carry genes that are implicated in the interception and redirection of host translation. These genes include the initiation factors IF1 and IF3, as well as ribosomal proteins S4, S1, S21 and L7/L12 (ribosomal proteins were only recently reported in phages¹⁷ (Fig. 3)). Both rpS1 and rpS21 are important for translation initiation in bacteria^{18–20}, making them likely to be useful for the hijacking of host ribosomes. Further analysis of rpS21 proteins revealed N-terminal extensions that were rich in basic and aromatic residues important for RNA binding. We predict that these phage ribosomal proteins substitute for host proteins¹⁷, and their extensions assist in competitive ribosome binding or preferential initiation of phage mRNAs.

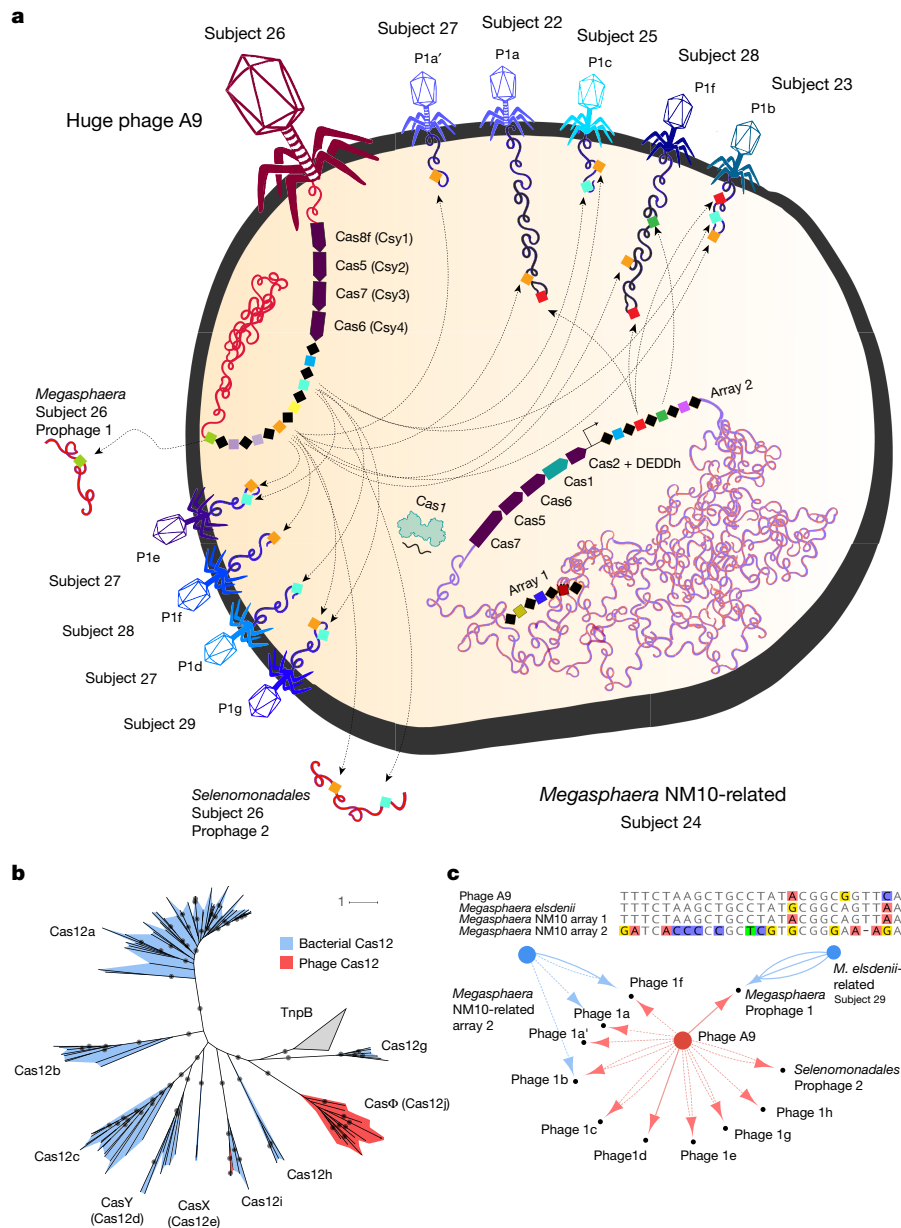


Fig. 4 | Phage and bacterial CRISPR-interaction dynamics. a, Cell diagram of bacterium–phage and phage–phage interactions that involve CRISPR targeting during superinfection. Arrows indicate CRISPR–Cas targeting of the prophage and phage genomes. Phage names indicate related groups delineated by whole-genome alignment. We only included CRISPR interactions from samples of subjects of the same human cohort. **b**, Maximum likelihood phylogenetic tree of Cas12 subtypes a–i. Phage-encoded Cas12i and CasΦ, the new effector, are outlined in red, with bacteria-encoded proteins in blue. Bootstrap values >90 are shown on the branches (circles). Cas14 and type V–U trees are provided

separately (Supplementary Fig. 11). Scale bars indicates the number of substitutions per site. **c**, Top, alignment of the consensus repeats from the A9 phage array and predicted host bacterial arrays. Bottom, interaction network showing the targeting of bacteria-encoded (blue) and phage-encoded (red) CRISPR spacers. The number of edges indicate the number of spacers from the array with targets to the smaller node. Solid edges denote spacer targets with no or one mismatch, and dashed edges denote two to three mismatches (to account for degeneration in old-end phage spacers, diversity in different subjects or phage mutation to avoid targeting).

Because rpS1 is often studied in the context of Shine–Dalgarno sequence recognition by the ribosome^{19,20}, we predicted the ribosomal binding sites for each phage genome (Methods). Whereas most phages have canonical Shine–Dalgarno sequences, huge phages from this study that carry possible rpS1s rarely have identifiable Shine–Dalgarno sequences (Supplementary Information and Supplementary Table 8). It is difficult to confirm ‘true’ rpS1 proteins owing to the ubiquity of the S1 domain, but this correlation with non-canonical Shine–Dalgarno sequences suggests a role in translation initiation, either on or off the ribosome.

Although assuming control of initiation may be the most logical step for the redirection of host translation by the phage, improving the efficiency of elongation and termination is necessary for robust infection and replication. Accordingly, we found many genes associated with the later steps of translation in phage genomes. These include elongation factors G, Tu and Ts, rPL7/12 and the processing enzyme peptide deformylase (Fig. 3), which has previously been reported in phage genomes²¹. We hypothesize that phage-encoded elongation factors maintain the overall translation efficiency during infection, much like the previously predicted role of peptide deformylase in

sustaining translation of the necessary photosynthetic proteins of the host²¹. Translation termination factors are also represented in our huge phage genomes, including release factor 1 and 2, ribosome recycling factor, as well as transfer messenger RNAs (tmRNAs) and small protein B (SmpB), which rescue ribosomes stalled on damaged transcripts and trigger the degradation of aberrant proteins. These tmRNAs are also used by phages to sense the physiological state of host cells and can induce lysis when the number of stalled ribosomes in the host is high²². Notably, some large putative plasmids have analogous suites of translation-relevant genes (Supplementary Table 5).

CRISPR–Cas-mediated interactions

We identified most major types of CRISPR–Cas systems in phages, including Cas9-based type II, the recently described type V-I²³, new variants of the type V-U systems²⁴ and new subtypes of the type V-F system²⁵ (Extended Data Fig. 8). The class II systems (types II and V) have not previously been reported in phages. Most phage effector nucleases (for interference) have conserved catalytic residues, implying that they are functional.

In contrast to the well-described case of a phage with a CRISPR system²⁶, almost all phage CRISPR systems lack spacer acquisition machinery (Cas1, Cas2 and Cas4) and many lack recognizable genes for interference (Extended Data Fig. 9 and Supplementary Table 1). For example, two related phages have a type I-C variant system that lacks Cas1 and Cas2 and have a helicase protein instead of Cas3. These phages also have a second system that contains a new candidate type V effector protein, CasΦ (Cas12j), which is approximately 750 amino acids in length (Fig. 4 and Supplementary Table 1), which occurs proximal to CRISPR arrays.

In some cases, phages that lack genes for interference and spacer integration have similar CRISPR repeats as their hosts (Fig. 4c) and may therefore use the Cas proteins of the host. Alternatively, systems that lack an effector nuclease may repress the transcription of the target sequences without cleavage^{27,28}. Additionally, spacer-repeat guide RNAs may have an RNA-interference-like mechanism to silence host CRISPR systems or nucleic acids to which they can hybridize. The phage-encoded CRISPR arrays are often compact (median, six repeats per array) (Extended Data Fig. 10). This range is substantially smaller than typically found in prokaryotic genomes (mean of 41 repeats for class I systems)²⁹. Some phage spacers target core structural and regulatory genes of other phages (Fig. 4c and Supplementary Table 10). Thus, phages apparently augment the immune arsenal of their hosts to prevent infection by competing phages.

Some phage-encoded CRISPR loci have spacers that target bacteria in the same sample or in a sample from the same study. We suppose that the targeted bacteria are the hosts for these phages, an inference supported by other host prediction analyses (Supplementary Table 4). Some loci with bacterial chromosome-targeting spacers encode Cas proteins that could cleave the host chromosome, whereas others do not. The targeting of host genes could disable or alter their regulation, which may be advantageous during the phage infection cycle. Some phage CRISPR spacers target bacterial intergenic regions, possibly interfering with genome regulation by blocking promoters or silencing non-coding RNAs.

Notable examples of CRISPR targeting of bacterial chromosomes involve transcription and translation genes. For instance, one phage targets a σ^{70} transcription factor gene in the genome of its host and encodes its own σ^{70} (Supplementary Information). Some huge phage genomes encode anti-sigma factor-like proteins (AsiA), consistent with previous reports of σ^{70} hijacking by phages with AsiA³⁰. In another example, a phage spacer targets the host glycyl tRNA synthetase, but the Cas14 effector lacks one of the required catalytic residues for cleavage, suggesting a role in repression (as a ‘dCas14’), rather than in cleavage (Supplementary Information).

Notably, we found no evidence of host-encoded spacers that target any CRISPR-bearing phages. However, phage CRISPR targeting of other

phages that are also targeted by bacterial CRISPR (Fig. 4c) suggested phage–host associations that were broadly confirmed by the phage taxonomic profile (Supplementary Table 4).

Some large *Pseudomonas*-infecting phages encode anti-CRISPRs^{31,32} (Acrs) and proteins that assemble a nucleus-like compartment that segregates their replicating genomes from host-defence and other bacterial systems³³. We identified proteins encoded in huge phage genomes that cluster with AcrVA5, AcrVA2, AcrIIA7 and AcrIIA11 and may function as Acrs. We also identified tubulin homologues (PhuZ) and proteins (Supplementary Information) that create a proteinaceous phage ‘nucleus’³⁴. The phage nucleus was recently shown to protect the phage genome against host defence by physically blocking degradation by CRISPR–Cas systems³⁵.

Conclusions

We show that phages with huge genomes are widespread across Earth’s ecosystems. We manually completed 35 genomes, distinguishing them from prophages, providing accurate genome lengths and complete inventories of genes, including those encoded in complex repeat regions that break automated assemblies. Even closely related phages have diversified across habitats. Host and phage migration could transfer genes relevant to medicine and agriculture (for example, genes that affect pathogenicity and antibiotic resistance) (Supplementary Information). Additional mechanisms that are relevant to medical applications involve the direct or indirect activation of immune responses. For example, some phages directly stimulate IFN γ through a TLR9-dependent pathway and exacerbate colitis³⁶. Huge phages may represent a reservoir of novel nucleic acid manipulation tools with applications in genome editing and might be harnessed to improve human and animal health. For instance, huge phages equipped with CRISPR–Cas systems might be tamed and used to modulate the functions of the bacterial microbiome or eliminate unwanted bacteria.

The huge phages comprise extensive clades, suggesting that a gene inventory comparable in size to those of many symbiotic bacteria is a conserved strategy for phage survival. Overall, their genes appear to redirect the protein production capacity of the host to favour phage genes by first intercepting the earliest steps of translation and subsequently ensuring the efficient production of proteins. These inferences are aligned with findings for some eukaryotic viruses, which control every phase of protein synthesis³⁷. Some phages acquired CRISPR–Cas systems with unusual compositions that may function to control host genes and eliminate competing phages.

More broadly, huge phages represent little-known biology, the platforms for which are distinct from those of small phages and partially analogous to those of symbiotic bacteria, blurring the distinctions between life and non-life. Given phylogenetic evidence for large radiations of huge phages, we wonder whether they are ancient and arose simultaneously with free-living cells, their symbionts and other phages from a pre-life (protogenote) state³⁸ rather than appearing more recently through episodes of genome expansion.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2007-4>.

1. Yuan, Y. & Gao, M. Jumbo bacteriophages: an overview. *Front. Microbiol.* **8**, 403 (2017).
2. Breitbart, M., Bonnain, C., Malki, K. & Sawaya, N. A. Phage puppet masters of the marine microbial realm. *Nat. Microbiol.* **3**, 754–766 (2018).
3. Rascovan, N., Duraisamy, R. & Desnues, C. Metagenomics and the human virome in asymptomatic individuals. *Annu. Rev. Microbiol.* **70**, 125–141 (2016).
4. Emerson, J. B. et al. Host-linked soil viral ecology along a permafrost thaw gradient. *Nat. Microbiol.* **3**, 870–880 (2018).

5. Balcazar, J. L. Bacteriophages as vehicles for antibiotic resistance genes in the environment. *PLoS Pathog.* **10**, e1004219 (2014).
6. Penadés, J. R., Chen, J., Quiles-Puchalt, N., Carpena, N. & Novick, R. P. Bacteriophage-mediated spread of bacterial virulence genes. *Curr. Opin. Microbiol.* **23**, 171–178 (2015).
7. Brown-Jaque, M. et al. Detection of bacteriophage particles containing antibiotic resistance genes in the sputum of cystic fibrosis patients. *Front. Microbiol.* **9**, 856 (2018).
8. Shkoporov, A. N. & Hill, C. Bacteriophages of the human gut: the “known unknown” of the microbiome. *Cell Host Microbe* **25**, 195–209 (2019).
9. Devoto, A. E. et al. Megaphages infect *Prevotella* and variants are widespread in gut microbiomes. *Nat. Microbiol.* **4**, 693–700 (2019).
10. Castelle, C. J. et al. Biosynthetic capacity, metabolic variety and unusual biology in the CPR and DPANN radiations. *Nat. Rev. Microbiol.* **16**, 629–645 (2018).
11. Pérez-Brocal, V. et al. A small microbial genome: the end of a long symbiotic relationship? *Science* **314**, 312–313 (2006).
12. Nakabachi, A. et al. The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science* **314**, 267 (2006).
13. Lobry, J. R. Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol. Biol. Evol.* **13**, 660–665 (1996).
14. Paez-Espino, D. et al. Uncovering Earth's virome. *Nature* **536**, 425–430 (2016).
15. Paez-Espino, D. et al. IMG/VR: a database of cultured and uncultured DNA viruses and retroviruses. *Nucleic Acids Res.* **45**, D457–D465 (2017).
16. Ivanova, N. N. et al. Stop codon reassignments in the wild. *Science* **344**, 909–913 (2014).
17. Mizuno, C. M. et al. Numerous cultivated and uncultivated viruses encode ribosomal proteins. *Nat. Commun.* **10**, 752 (2019).
18. van Duin, J. & Wijnands, R. The function of ribosomal protein S21 in protein synthesis. *Eur. J. Biochem.* **118**, 615–619 (1981).
19. Farwell, M. A., Roberts, M. W. & Rabinowitz, J. C. The effect of ribosomal protein S1 from *Escherichia coli* and *Micrococcus luteus* on protein synthesis in vitro by *E. coli* and *Bacillus subtilis*. *Mol. Microbiol.* **6**, 3375–3383 (1992).
20. Sørensen, M. A., Fricke, J. & Pedersen, S. Ribosomal protein S1 is required for translation of most, if not all, natural mRNAs in *Escherichia coli* in vivo. *J. Mol. Biol.* **280**, 561–569 (1998).
21. Frank, J. A. et al. Structure and function of a cyanophage-encoded peptide deformylase. *ISME J.* **7**, 1150–1160 (2013).
22. Janssen, B. D. & Hayes, C. S. The tmRNA ribosome-rescue system. *Adv. Protein Chem. Struct. Biol.* **86**, 151–191 (2012).
23. Yan, W. X. et al. Functionally diverse type V CRISPR–Cas systems. *Science* **363**, 88–91 (2019).
24. Shmakov, S. et al. Diversity and evolution of class 2 CRISPR–Cas systems. *Nat. Rev. Microbiol.* **15**, 169–182 (2017).
25. Harrington, L. B. et al. Programmed DNA destruction by miniature CRISPR–Cas14 enzymes. *Science* **362**, 839–842 (2018).
26. Seed, K. D., Lazinski, D. W., Calderwood, S. B. & Camilli, A. A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. *Nature* **494**, 489–491 (2013).
27. Luo, M. L., Mullis, A. S., Leenay, R. T. & Beisel, C. L. Repurposing endogenous type I CRISPR–Cas systems for programmable gene repression. *Nucleic Acids Res.* **43**, 674–681 (2015).
28. Stachler, A.-E. & Marchfelder, A. Gene repression in Haloarchaea using the CRISPR (clustered regularly interspaced short palindromic repeats)–Cas I–B system. *J. Biol. Chem.* **291**, 15226–15242 (2016).
29. Toms, A. & Barrangou, R. On the global CRISPR array behavior in class I systems. *Biol. Direct* **12**, 20 (2017).
30. Brown, K. L. & Hughes, K. T. The role of anti-sigma factors in gene regulation. *Mol. Microbiol.* **16**, 397–404 (1995).
31. Bondy-Denomy, J. et al. Multiple mechanisms for CRISPR–Cas inhibition by anti-CRISPR proteins. *Nature* **526**, 136–139 (2015).
32. Pawluk, A. et al. Inactivation of CRISPR–Cas systems by anti-CRISPR proteins in diverse bacterial species. *Nat. Microbiol.* **1**, 16085 (2016).
33. Chaikereetisak, V. et al. Assembly of a nucleus-like structure during viral replication in bacteria. *Science* **355**, 194–197 (2017).
34. Chaikereetisak, V. et al. The phage nucleus and tubulin spindle are conserved among large *Pseudomonas* phages. *Cell Rep.* **20**, 1563–1571 (2017).
35. Mendoza, S. D. et al. A bacteriophage nucleus-like compartment shields DNA from CRISPR nucleases. *Nature* **577**, 244–248 (2020).
36. Gogokhia, L. et al. Expansion of bacteriophages is linked to aggravated intestinal inflammation and colitis. *Cell Host Microbe* **25**, 285–299 (2019).
37. Jaafar, Z. A. & Kieft, J. S. Viral RNA structure-based strategies to manipulate translation. *Nat. Rev. Microbiol.* **17**, 110–123 (2019).
38. Woese, C. The universal ancestor. *Proc. Natl Acad. Sci. USA* **95**, 6854–6859 (1998).
39. Subramanian, A. R. Structure and functions of ribosomal protein S1. *Prog. Nucleic Acid Res. Mol. Biol.* **28**, 101–142 (1983).
40. Loveland, A. B. & Korostelev, A. A. Structural dynamics of protein S1 on the 70S ribosome visualized by ensemble cryo-EM. *Methods* **137**, 55–66 (2018).
41. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

Methods

Phage- and plasmid-genome identification

Datasets generated in the current study, those from previous research conducted by our team, the *Tara* Oceans microbiomes⁴² and the Global Oceans Virome⁴³ were searched for sequence assemblies that could have derived from phages with genomes of more than 200 kb in length. Read assembly, gene prediction and initial gene annotation followed standard, previously reported methods^{44–48}.

Phage candidates were initially found by retrieving sequences that were not assigned to a genome and had no clear taxonomic profile at the domain level. Taxonomic profiles were determined through a voting scheme, in which the winning taxonomy had to have more than 50% votes for each taxonomic rank on the basis of protein annotations in the UniProt and ggKbase (<https://ggkbase.berkeley.edu/>) databases⁴⁹. Phages were further narrowed down by identifying sequences with a high number of hypothetical protein annotations and/or the presence of phage-specific genes, such as capsid, tail, terminase, spike, holin, portal and baseplate. All candidate phage sequences were checked throughout to distinguish putative prophages from phages. Prophages were identified on the basis of a clear transition into the host genome with a high fraction of confident functional predictions, often associated with core metabolic functions and much higher similarity to bacterial genomes. Plasmids were distinguished from phages on the basis of matches to plasmid partitioning and conjugative transfer genes. Those that did not have phage-specific genes were assigned using phylogenetic tree placement using *recA*, *polA*, *polB*, *dnaE* and the DNA sliding clamp loader gene. Phages and placement assignments were further verified using a network of protein clustering with proteins from RefSeq prokaryotic viruses and 400 randomly sampled plasmids of more than 200 kb using vContact2⁵⁰ (Extended Data Fig. 6).

Phage- and plasmid-genome manual curation

All classified scaffolds were tested for end overlaps indicative of circularization. Assembled sequences that could be perfectly circularized were considered potentially complete. Erroneous concatenated sequence assemblies were initially flagged by searching for direct repeats of more than 5 kb using Vmatch⁵¹. Potentially concatenated sequence assemblies were manually checked for multiple large repeating sequences using the dotplot and RepeatFinder features in Geneious v.9. Sequences were corrected and removed from further analysis if the corrected length was more than 200 kb.

A subset of the phage sequences was selected for manual curation, with the goal of finishing (replacing all Ns at scaffolding gaps or local misassemblies by the correct nucleotide sequences and circularization). Curation generally followed previously described methods⁹. In brief, reads from the appropriate dataset were mapped using Bowtie2 v.2.3.4.1⁵² to the de novo assembled sequences. Unplaced mate pairs of mapped reads were retained with shrinksam (<https://github.com/bcthomas/shrinksam>). Mappings were manually checked throughout to identify local misassemblies using Geneious v.9. N-filled gaps or misassembly corrections made use of unplaced paired reads, in some cases using reads relocated from sites to which they were mismapped. In such cases, mismappings were identified on the basis of much larger than expected paired read distances, high polymorphism densities, backwards mapping of one read pair or any combination of these. Similarly, ends were extended using unplaced or incorrectly placed paired reads until circularization could be established. In some cases, extended ends were used to recruit new scaffolds that were then added to the assembly. The accuracy of all extensions and local assembly changes were verified in a subsequent phase of read mapping. In many cases, assemblies were terminated or internally corrupted by the presence of repeated sequences. In these cases, blocks of repeated sequences as well as unique flanking sequences were identified. Reads were then manually relocated, respecting paired-read placement rules and unique

flanking sequences. After gap closure, circularization and verification of accuracy throughout, end overlap was eliminated, genes were predicted and the start moved to an intergenic region, which was—in some cases—suspected to be origin on the basis of a combination of coverage trends and GC skew⁵³. Finally, the sequences were checked to identify any repeated sequences that could have led to an incorrect path choice because the repeated regions were larger than the distance spanned by paired reads. This step also ruled out artefactual long phage sequences generated by end-to-end repeats of smaller phages, which occur in previously described datasets⁹.

Structural and functional annotations

After the identification and curation of phage genomes, coding sequences and Shine–Dalgarno ribosomal binding site motifs were predicted with Prodigal using genetic code 11 (–m –g 11 –p single). The resulting coding sequences were annotated as previously described by searching UniProt, UniRef100 and KEGG⁵⁴. Functional annotations were further assigned by searching proteins in PFAM r32⁵⁵, TIGRFAMS r15⁵⁶, Virus Orthologous Groups (VOG) r90 (<http://vogdb.org/>) and Prokaryotic Virus Orthologous Groups⁵⁷ (pVOG). tRNAs were identified with tRNAscan-s.e. v.2.0⁵⁸ using the bacterial model. tmRNAs were assigned using ARAGORN v.1.2.38⁵⁹ with the genetic code of bacteria and plant chloroplasts.

Clustering of the coding sequences into families was achieved using a two-step procedure. A first protein clustering was done using the fast and sensitive protein-sequence searching software MMseqs⁶⁰. An all-versus-all sequences search was performed using an *E*-value cut-off of 1×10^{-3} , sensitivity of 7.5 and coverage of 0.5. A sequence similarity network was built on the basis of the pairwise similarities and the greedy set cover algorithm from MMseqs was performed to define protein subclusters. The resulting subclusters were defined as subfamilies. To test for distant homology, we grouped subfamilies into protein families using a comparison of hidden Markov models (HMMs). The proteins of each subfamily with at least two protein members were aligned using the result2msa parameter of MMseqs, and HMM profiles were built using the HHpred⁶¹ suite from the multiple sequence alignments. The subfamilies were then compared to each other using HHblits from the HHpred suite (with parameters –v 0 –p 50 –z 4 –Z 32000 –B 0 –b 0). For subfamilies with probability scores of at least 95% and coverage at least 0.50, a similarity score (probability \times coverage) was used as weight of the input network in the final clustering using the Markov clustering algorithm⁶², with 2.0 as the inflation parameter. These clusters were defined as the protein families. Protein sequences were functionally annotated on the basis of their best hmmsearch match (v.3.1) (*E*-value cut-off 1×10^{-3}) against an HMM database constructed on the basis of orthologous groups defined by the KEGG database⁶³ (downloaded on 10 June 2015). Domains were predicted using the same hmmsearch procedure against the PFAM r31 database⁵⁵. The domain architecture of each protein sequence was predicted using the DAMA software⁶⁴ (default parameters). SIGNALP⁶⁵ (v.4.1) (parameters, –f short –t gram+) and PSORT⁶⁶ v.3 (parameters, –long –positive) were used to predict the putative cellular localization of the proteins. Prediction of transmembrane helices in proteins was performed using TMHMM⁶⁷ (v.2.0) (default parameters). Hairpins (palindromes, based on identical overlapping repeats in the forward and reverse directions) were identified using the Geneious Repeat Finder and located across the dataset using Vmatch⁵¹. Repeats of more than 25 bp with 100% similarity were tabulated.

Reference genomes for size comparisons

RefSeq r92 genomes were recovered using the NCBI Virus portal and selecting only complete dsDNA genomes with bacterial hosts. Genomes from a previously published study¹⁴ were downloaded from IMG/VR and only sequence assemblies that were labelled ‘circular’ with predicted bacterial hosts were retained. Given the presence of sequences in IMG/

VR that were based on erroneous concatenations, we only considered sequences from this source that were more than 200 kb; however, a subset of these was removed as artefactual sequences.

Alternative genetic codes

In cases in which the gene prediction using the standard bacterial code (code 11) resulted in seemingly anomalously low coding densities, potential alternative genetic codes were investigated. In addition to making a prediction using the fast and accurate genetic code inference and logo⁶⁸ (FACIL) web server, we identified genes with well-defined functions (for example, polymerase or nuclease) and determined the stop codons terminating genes that were shorter than expected. We then repredicted genes using GLIMMER3 v.1.5⁶⁹ and Prodigal with TAG not interpreted as a stop codon. Other combinations of repurposed stop codons were evaluated and candidate codes (for example, code 6, with only one stop codon) were ruled out owing to unlikely gene-fusion predictions.

Large terminase subunit and MCP phylogenetic analyses

The phylogenetic tree of the large terminase subunit was constructed by recovering large terminases from the aforementioned protein-clustering and annotation pipeline. The coding sequences that matched with >30 bitscore to PFAM, TIGRFAMS, VOG and pVOG were retained. Any coding sequence that had a hit to large terminase, regardless of bitscore, was searched using HHblits⁷⁰ against the unclust30_2018_08 database. The resulting alignment was then further searched against the PDB70 database. Remaining coding sequences that clustered in protein families with a large terminase HMM were also included after manual verification. Detected large terminases were manually verified using the HHPred⁷⁰ and jPred⁷¹ web servers. Large terminases from the >200-kb phage genomes¹⁴ and all >200-kb complete dsDNA phage genomes from RefSeq r92 were also included by protein family clustering with the phage-coding sequences from this study. The resulting terminases were clustered at 95% amino acid identity to reduce redundancy using CD-HIT⁷². Smaller phage genomes were included by searching the resulting coding sequences set against the full RefSeq protein database and retaining the top 10 best hits. Those hits that had no large terminase match against PFAM, TIGRFAMS, VOG or pVOG were removed from further consideration and the remaining set was clustered at 90% amino acid identity. The final set of large terminase coding sequences that were more than 100 amino acids in length were aligned using MAFFT⁷³ v.7.407 (--localpair --maxiterate 1000) and poorly aligned sequences were removed and the resulting set was realigned. The phylogenetic tree was inferred using IQTREE v.1.6.6 using automatic model selection⁷⁴. The phylogenetic tree of MCP genes was constructed by retrieving all MCPs annotated by combining the PFAM annotations of protein families and direct annotations by PFAM, TIGRFAMS, VOG and pVOG. Reference MCP gene sequences were collected using the same strategy and sources as for the large terminase subunit tree. The resulting set was further screened by searching against PFAM, TIGRFAMS, VOG and pVOG and removing matches that had no large terminase match regardless of bitscore. The final set of MCP sequences were aligned with MAFFT(--localpair --maxiterate 1000) and the phylogenetic tree was constructed using IQTREE with automatic model selection and 1,000 bootstrap replicates.

Whole-genome scale clustering

To identify phage genomes that were closely related at the whole-genome level, we compared sequences using whole-genome alignments. The goal of this analysis was to further corroborate the identified phylogenetic clades and test for the presence of very similar phages in different habitats and environments. Genomes grouped together in the primary clusters from dRep v.2⁷⁵ were evaluated for genome alignment using Mauve⁷⁶ within Geneious v.9.

CRISPR–Cas locus and target detection

Phage- and host-encoded CRISPR loci (repeats and spacers) were identified using a combination of MinCED (<https://github.com/ctSkennerton/minced>) and CRISPRDetect⁷⁷. A custom database of Cas genes was built by collecting Cas gene sequences from previous studies^{23,25,78–82} and built with MAFFT (--localpair --maxiterate 1000) and hmmbuild. The coding sequences from this study were searched against the HMM database using hmmsearch with $E < 1 \times 10^{-5}$. Matches were checked using a combination of hmmscan and BLAST searches against the NCBI nr database and manually verified by identifying collocated CRISPR arrays and Cas genes. Spacers extracted from between repeats of the CRISPR locus were compared to sequences assemblies from the same site using BLASTN-short⁸³. Matches with alignment length >24 bp and ≤1 mismatch were retained and targets were classified as bacteria, phage or other. CRISPR arrays that had ≤1 mismatch, were further searched for more spacer matches in the target sequence by finding more hits with ≤3 mismatches.

Host identification

The phylum affiliations of bacterial hosts for phage and plasmid-like sequences were predicted by considering the UniProt taxonomic profiles of every coding sequence for each phage genome. The phylum level matches for each phage genome were summed and the phylum with the most hits was considered as the potential host phylum. However, only cases in which this phylum that had 3× as many counts as the next most-counted phylum were assigned as the tentative phage host phylum. Phage hosts were further assigned and verified using the CRISPR-targeting strategy describe above with the phage and plasmid-like genomes as targets. CRISPR arrays were predicted on all sequence assemblies from the same site that each phage genome was reconstructed. Sequence assemblies containing spacers with a match of length >24 bp and ≤1 mismatch were used to infer phage–host relationships. In all cases, the predicted host phylum based on taxonomic profiling and CRISPR targeting were in complete agreement. Similarly, the phyla of hosts were predicted on the basis of phylogenetic analysis of phage genes also found in host genomes (for example, involved in translation and nucleotide reactions). Inferences based on computed taxonomic profiles and phylogenetic trees were also in complete agreement.

Phage-encoded tRNA synthetase trees

Phylogenetic trees were constructed for phage-encoded tRNA synthetase, ribosomal and initiation factor protein sequences using a set of the closest reference sequences from NCBI and bacterial genomes from the current study. The tRNA synthetases were identified on the basis of annotation of genes using the standard ggKbase pipeline (see above), and confirmed by HMMs with datasets from TIGRFAMS. For each type of tRNA synthetase, references were selected by comparing all of the corresponding genes of this type against the NCBI nr database using DIAMOND v.0.9.24⁸⁴, their top 100 hits were clustered by CD-HIT using a 90% similarity threshold⁷². The phylogenetic tree of each tRNA synthetase was constructed using RAxML v.8.0.26⁸⁵ with the PROTGAMMALG model.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

GenBank files for all genomes are provided as Supplementary Information. Sequence reads and genomes have been deposited at the European Nucleotide Archive (ENA) under project PRJEB35371. Genomes have been deposited at ENA under accessions ERS4026114–ERS4026474.

Reads are available at ENA under accessions ERS4025670–ERS4025731. Read accessions and genome accessions for each phage genome are included in Supplementary Table 1.

Code availability

The custom code used to analyse the genomes is available at http://www.github.com/rohansachdeva/assembly_repeats.

42. Karsenti, E. et al. A holistic approach to marine eco-systems biology. *PLoS Biol.* **9**, e1001177 (2011).
43. Roux, S. et al. Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature* **537**, 689–693 (2016).
44. Peng, Y., Leung, H. C. M., Yiu, S. M. & Chin, F. Y. L. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* **28**, 1420–1428 (2012).
45. Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* **27**, 824–834 (2017).
46. Edgar, R. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
47. Joshi, N. A. & Fass, J. N. Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files. v1.33 <https://github.com/najoshi/sickle> (2011).
48. Bushnell, B. BMap short read aligner. <https://sourceforge.net/projects/bbmap/> (2019).
49. Raveh-Sadka, T. et al. Gut bacteria are rarely shared by co-hospitalized premature infants, regardless of necrotizing enterocolitis development. *eLife* **4**, e05477 (2015).
50. Bulduc, B. et al. vCONTACT: an iVirus tool to classify double-stranded DNA viruses that infect Archaea and Bacteria. *PeerJ* **5**, e3243 (2017).
51. Kurtz, S. The Vmatch large scale sequence analysis software. <http://www.vmatch.de/> (2017).
52. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
53. Brown, C. T., Olm, M. R., Thomas, B. C. & Banfield, J. F. Measurement of bacterial replication rates in microbial communities. *Nat. Biotechnol.* **34**, 1256–1263 (2016).
54. Wrighton, K. C. et al. Metabolic interdependencies between phylogenetically novel fermenters and respiratory organisms in an unconfined aquifer. *ISME J.* **8**, 1452–1463 (2014).
55. Finn, R. D. et al. Pfam: the protein families database. *Nucleic Acids Res.* **42**, D222–D230 (2014).
56. Haft, D. H. et al. TIGRFAMs and genome properties in 2013. *Nucleic Acids Res.* **41**, D387–D395 (2013).
57. Graziotin, A. L., Koonin, E. V. & Kristensen, D. M. Prokaryotic Virus Orthologous Groups (pVOGs): a resource for comparative genomics and protein family annotation. *Nucleic Acids Res.* **45**, D491–D498 (2017).
58. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–964 (1997).
59. Laslett, D. & Canback, B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* **32**, 11–16 (2004).
60. Hauser, M., Steinegger, M. & Söding, J. MMseqs software suite for fast and deep clustering and searching of large protein sequence sets. *Bioinformatics* **32**, 1323–1330 (2016).
61. Remmert, M., Biegert, A., Hauser, A. & Söding, J. HHblits: lightning-fast iterative protein sequence searching by HMM–HMM alignment. *Nat. Methods* **9**, 173–175 (2012).
62. Enright, A. J., Van Dongen, S. & Ouzounis, C. A. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* **30**, 1575–1584 (2002).
63. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–D462 (2016).
64. Bernardes, J. S., Vieira, F. R. J., Zaverucha, G. & Carbone, A. A multi-objective optimization approach accurately resolves protein domain architectures. *Bioinformatics* **32**, 345–353 (2016).
65. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).
66. Peabody, M. A., Laird, M. R., Vlasschaert, C., Lo, R. & Brinkman, F. S. L. PSORTdb: expanding the bacteria and archaea protein subcellular localization database to better reflect diversity in cell envelope structures. *Nucleic Acids Res.* **44**, D663–D668 (2016).
67. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580 (2001).
68. Dutilh, B. E. et al. FACIL: fast and accurate genetic code inference and logo. *Bioinformatics* **27**, 1929–1933 (2011).
69. Delcher, A. L., Harmon, D., Kasif, S., White, O. & Salzberg, S. L. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* **27**, 4636–4641 (1999).
70. Steinegger, M., Meier, A. & Biegert, A. HH-suite3 for fast remote homology detection and deep protein annotation. *Bioinformatics* **30**, 473 (2019).
71. Cole, C., Barber, J. D. & Barton, G. J. The Jpred 3 secondary structure prediction server. *Nucleic Acids Res.* **36**, W197–W201 (2008).
72. Huang, Y., Niu, B., Gao, Y., Fu, L. & Li, W. CD-HIT suite: a web server for clustering and comparing biological sequences. *Bioinformatics* **26**, 680–682 (2010).
73. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

74. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
75. Olm, M. R., Brown, C. T., Brooks, B. & Banfield, J. F. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* **11**, 2864–2868 (2017).
76. Darling, A. C. E., Mau, B., Blattner, F. R. & Perna, N. T. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* **14**, 1394–1403 (2004).
77. Biswas, A., Staals, R. H. J., Morales, S. E., Fineran, P. C. & Brown, C. M. CRISPRDetect: a flexible algorithm to define CRISPR arrays. *BMC Genomics* **17**, 356 (2016).
78. Burstein, D. et al. New CRISPR–Cas systems from uncultivated microbes. *Nature* **542**, 237–241 (2017).
79. Makarova, K. S. et al. An updated evolutionary classification of CRISPR–Cas systems. *Nat. Rev. Microbiol.* **13**, 722–736 (2015).
80. Shmakov, S. et al. Discovery and functional characterization of diverse class 2 CRISPR–Cas systems. *Mol. Cell* **60**, 385–397 (2015).
81. Yan, W. X. et al. Cas13d is a compact RNA-targeting type VI CRISPR effector positively modulated by a WYL-domain-containing accessory protein. *Mol. Cell* **70**, 327–339 (2018).
82. Smargon, A. A. et al. Cas13b is a Type VI-B CRISPR-associated RNA-guided RNase differentially regulated by accessory proteins Csx27 and Csx28. *Mol. Cell* **65**, 618–630 (2017).
83. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
84. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).
85. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
86. Smoot, M. E., Ono, K., Ruschinski, J., Wang, P.-L. & Ideker, T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27**, 431–432 (2011).

Acknowledgements Funding for this project was provided by the National Institutes of Health (NIH) under awards RAI092531A and R01-GM109454 and the Alfred P. Sloan Foundation under grant APSF-2012-10-05 to J.F.B. and M.M.; National Science Foundation (NSF) Sustainable Chemistry grant 1349278 to J.F.B., NSF Graduate Research Fellowships to B.A.-S. (DGE 1752814) and M.R.O. (DGE 1106400); the Paul Allen Foundation Frontiers Group; Chan Zuckerberg Biohub; Innovative Genomics Institute; The Novo Nordisk Foundation (NNF16OC0021856) to P.M.; NASA 13R-0043 and CA AES to K.R.A. and K.F.; German Science Foundation postdoctoral scholarship to A.J.P. (DFG PR 1603/1-1); Camille & Henry Dreyfus Postdoctoral Fellowship in Environmental Chemistry to C.H.; Watershed Function Scientific Focus Area funded by the US Department of Energy, Office of Science, Office of Biological and Environmental Research (DE-AC02-05CH11231) to A.L., P.M.-C. and A.T.; NSF grants GRT00048468 and 1342701 to K.C.W. and M.B., and CHE-1740549 to J.H.D.C.; the Ministry of Economy, Trade and Industry of Japan to Y.A.; the March of Dimes Prematurity Research Center at Stanford University School of Medicine, the Thomas C. and Joan M. Merigan Endowment at Stanford University to D.A.R.; the National Research Foundation of South Africa (UID 64877) to S.T.L.H. and NSF CZO funding. We acknowledge the scientists who generated the public database sequences and thank J. Tung, E. Archie, F. Aarestrup and R. Kruger for contributing data and P. Pausch and G. Knott for discussions. The phage icon used in the graphics was created by Two Photon Art and the Innovative Genomics Institute.

Author contributions Analyses were conducted primarily by B.A.-S., R.S., L.-X.C., F.W. and J.F.B. Specifically, phylogenetic and gene inventory analyses were performed by B.A.-S. and R.S., tRNA synthetase analysis was led by L.-X.C., ribosome analyses were led by F.W. and B.A.-S. with input from J.H.D.C. and genome analysis scripts were written by R.S. R.M. provided code for protein clustering. Size-distribution and tRNA analysis was conducted by R.S. Genome curation was performed by J.F.B., P.M., L.-X.C., A.D. and C.S. B.A.-S. led the CRISPR–Cas analyses. J.A.D. provided support and scientific input. S.L. provided bioinformatics support. L.-X.C., P.M., A.D., C.J.C., M.R.O., K.B.-G., Y.A., C.H., B.B., A.T., A.L., P.M.-C., C.S., D.S.A.G., M.A.B., A.S., A.L.J., T.C.N., R. Kantor, R. Keren, K.R.L., I.F.F., K.F., R.A., K.A., J.Z., A.J.P., M.E.P., S.G.T., W.-J.L., K.W., S.H., M.M., D.A.R., A.-C.L., L.W. and J.M.S. contributed data. B.A.-S., R.S. and J.F.B. wrote the manuscript, with input from all authors. The study was conceived by J.F.B., with input from P.M. and A.D.

Competing interests The Regents of the University of California have patents pending for CRISPR technologies on which the authors are inventors. J.A.D. is a co-founder of Caribou Biosciences, Editas Medicine, Intellia Therapeutics, Scribe Therapeutics and Mammoth Biosciences, a scientific advisory board member of Caribou Biosciences, Intellia Therapeutics, eFFECTOR Therapeutics, Scribe Therapeutics, Synthego, Mammoth Biosciences and Inari, and is a Director at Johnson & Johnson and has sponsored research projects by Biogen and Pfizer. J.F.B. is a founder of Metagenomi.

Additional information

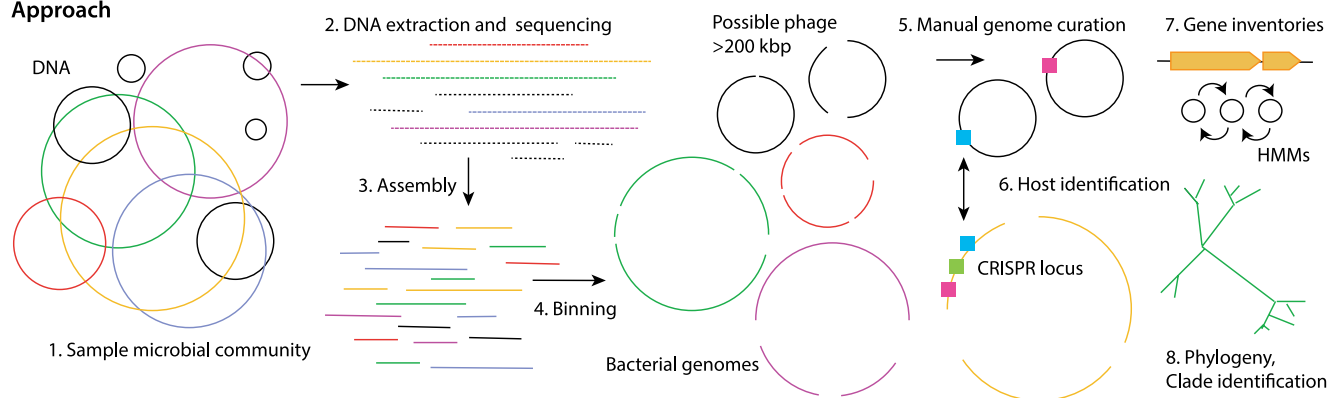
Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-2007-4>.

Correspondence and requests for materials should be addressed to J.F.B.

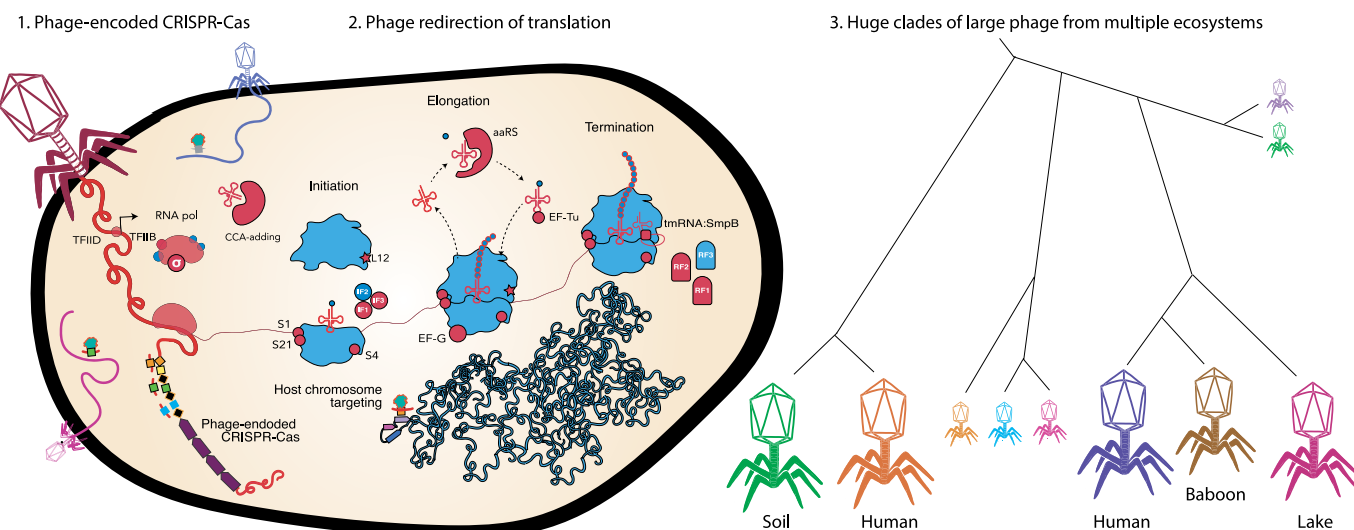
Peer review information *Nature* thanks Robert Edwards and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

Approach

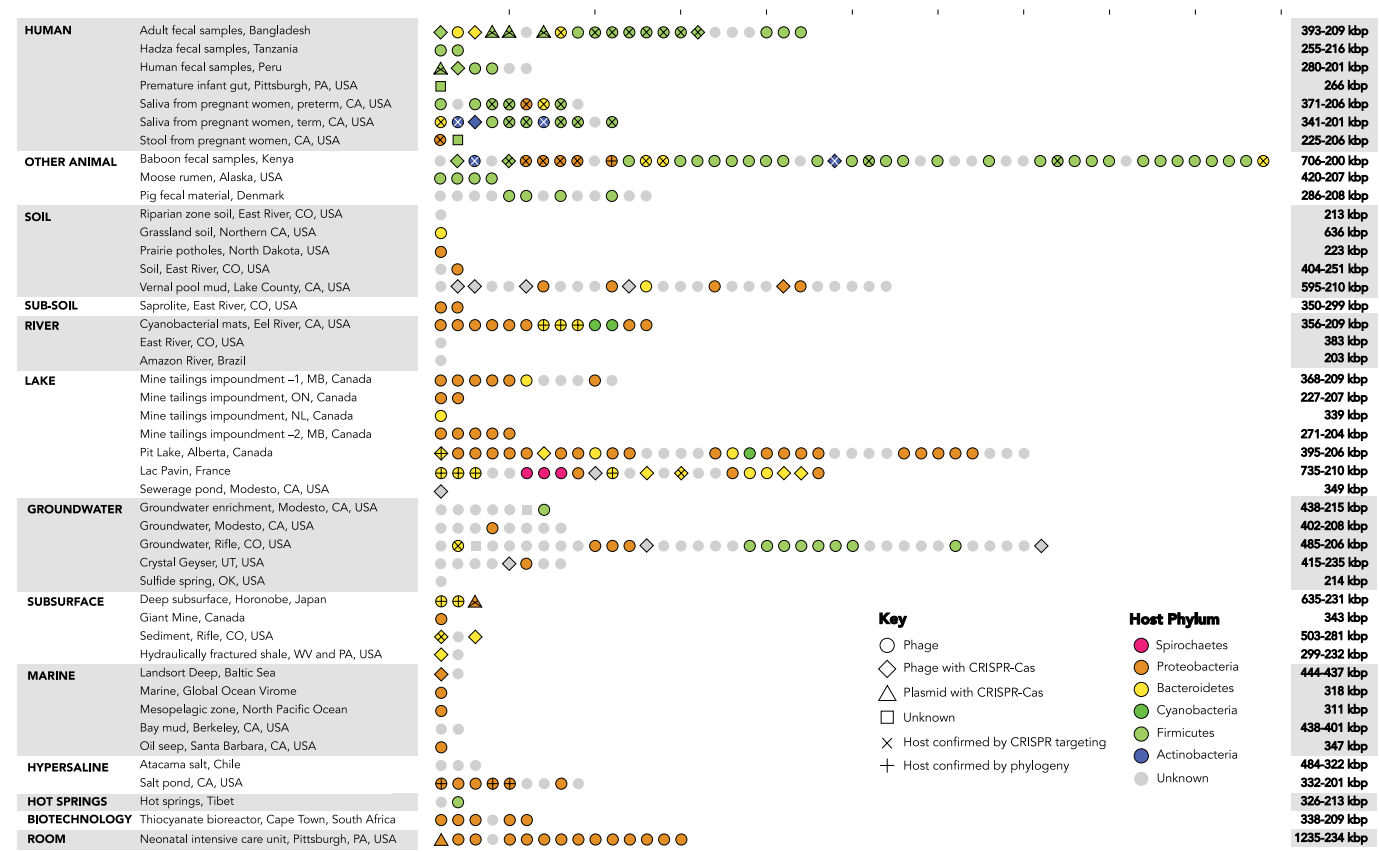


Main findings



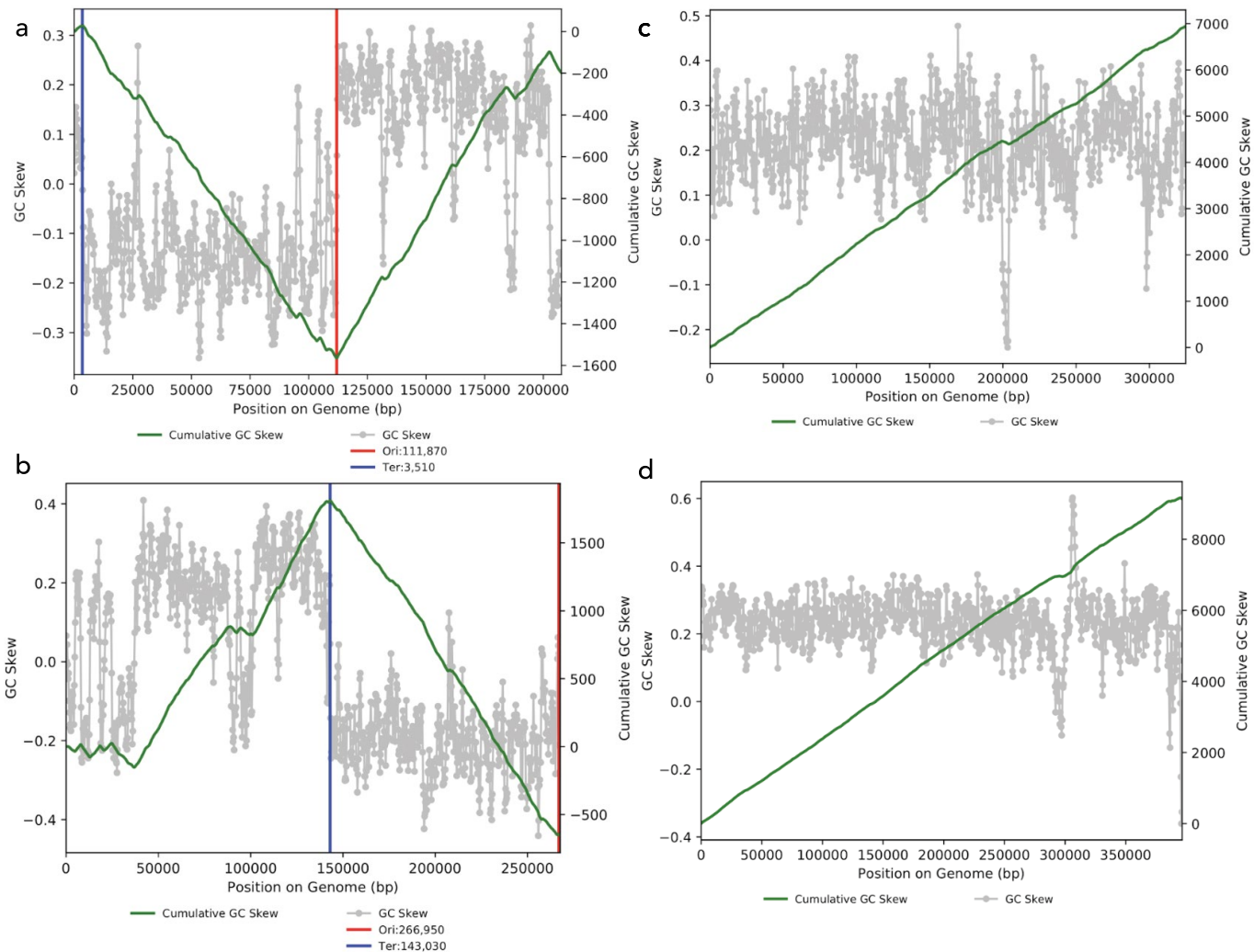
Extended Data Fig. 1 | Graphical abstract describing the approach and main findings of this study. aaRS, aminoacyl-tRNA synthetase; CCA-adding, tRNA nucleotidyltransferase; EF, elongation factor; IF, initiation factor; PDF, peptide

deformylase; QueC/D/F, queuosine synthesis and tRNA modification; RF, release factor; RNA Pol, RNA polymerase; RRF, ribosome recycling factor; TFIIB, transcription factor II B.



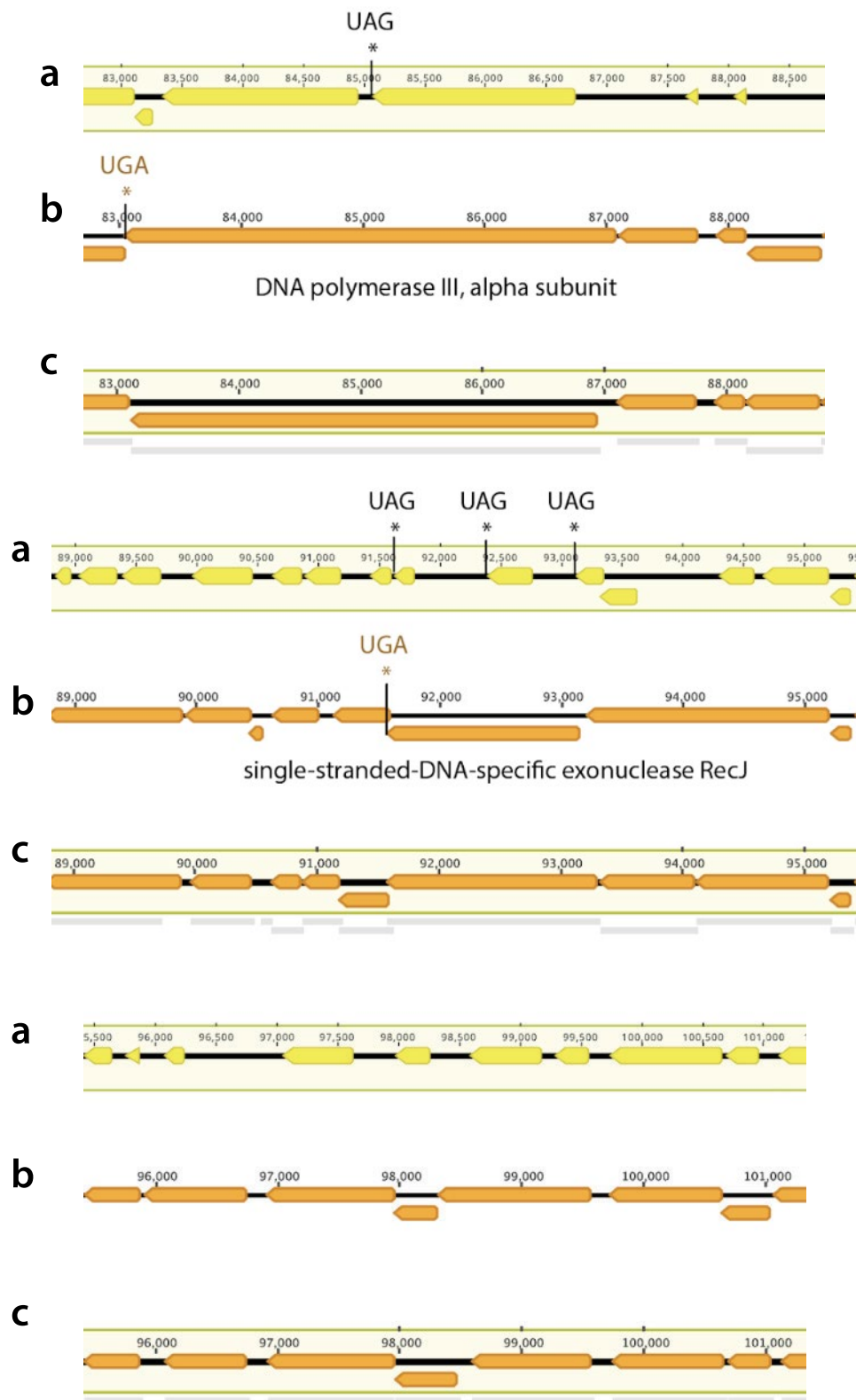
Extended Data Fig. 2 | Ecosystems with phage genomes and plasmid-like sequences of more than 200 kb. Genomes grouped by sampling-site type. Each box represents a phage genome or plasmid-like sequence, and boxes are horizontally arranged in order of decreasing genome size. The size range for

each site type is listed to the right. Colours indicate putative host phylum on the basis of genome taxonomic profile, with confirmation by CRISPR spacer targeting (X) or rps21 (+).



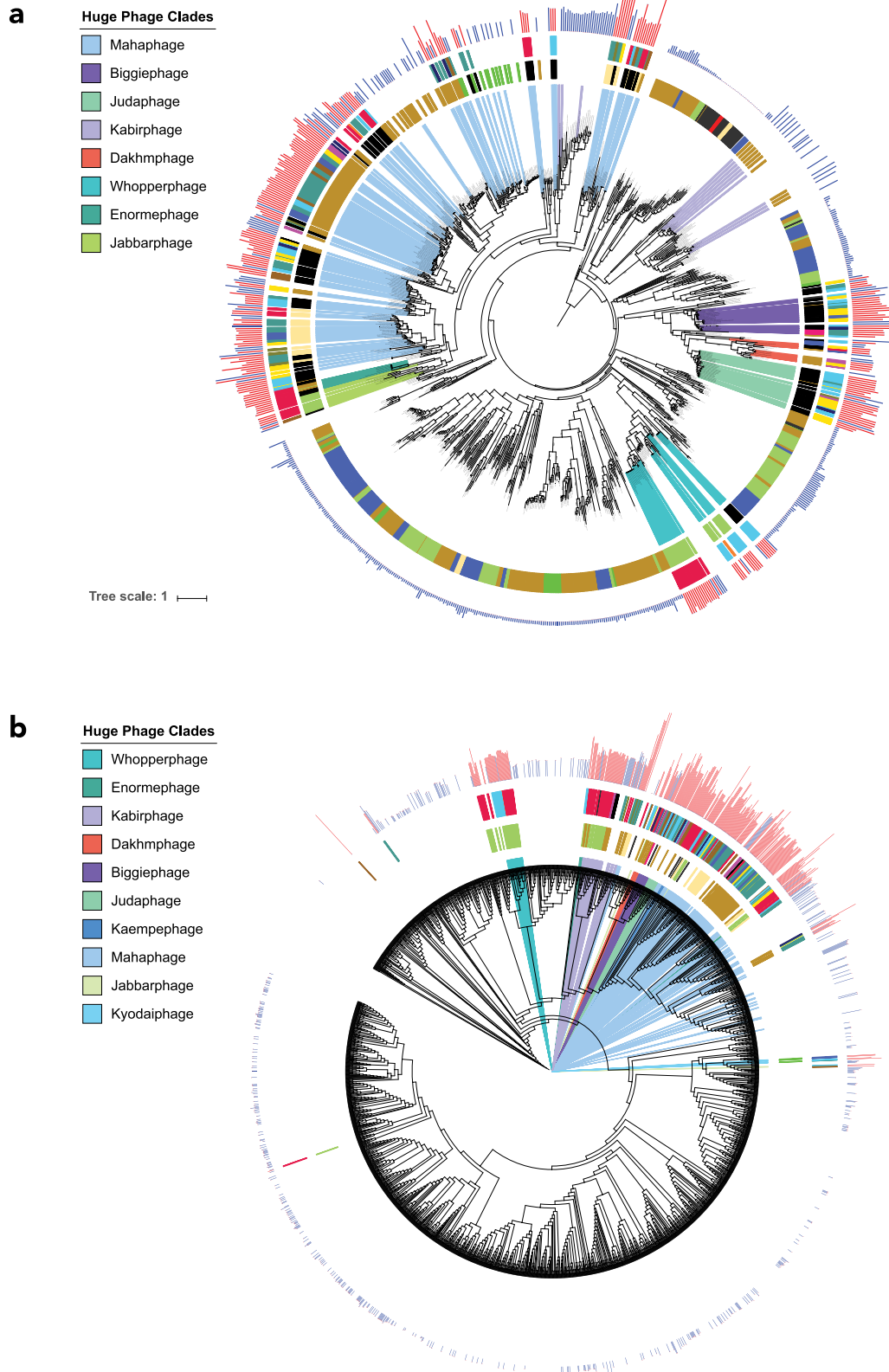
Extended Data Fig. 3 | Examples of phage genomes that display GC skew indicative of bidirectional replication. a, b, Example phage genomes with GC skew patterns that are strongly indicative of bidirectional replication (origin-

to-terminus) that is typically found in bacteria (however, the origin may not correspond to the start of the genome). **c, d,** Phage genomes with GC skew patterns that are suggestive of unidirectional patterns.

**Extended Data Fig. 4 | Example of the alternative coding of phages.**

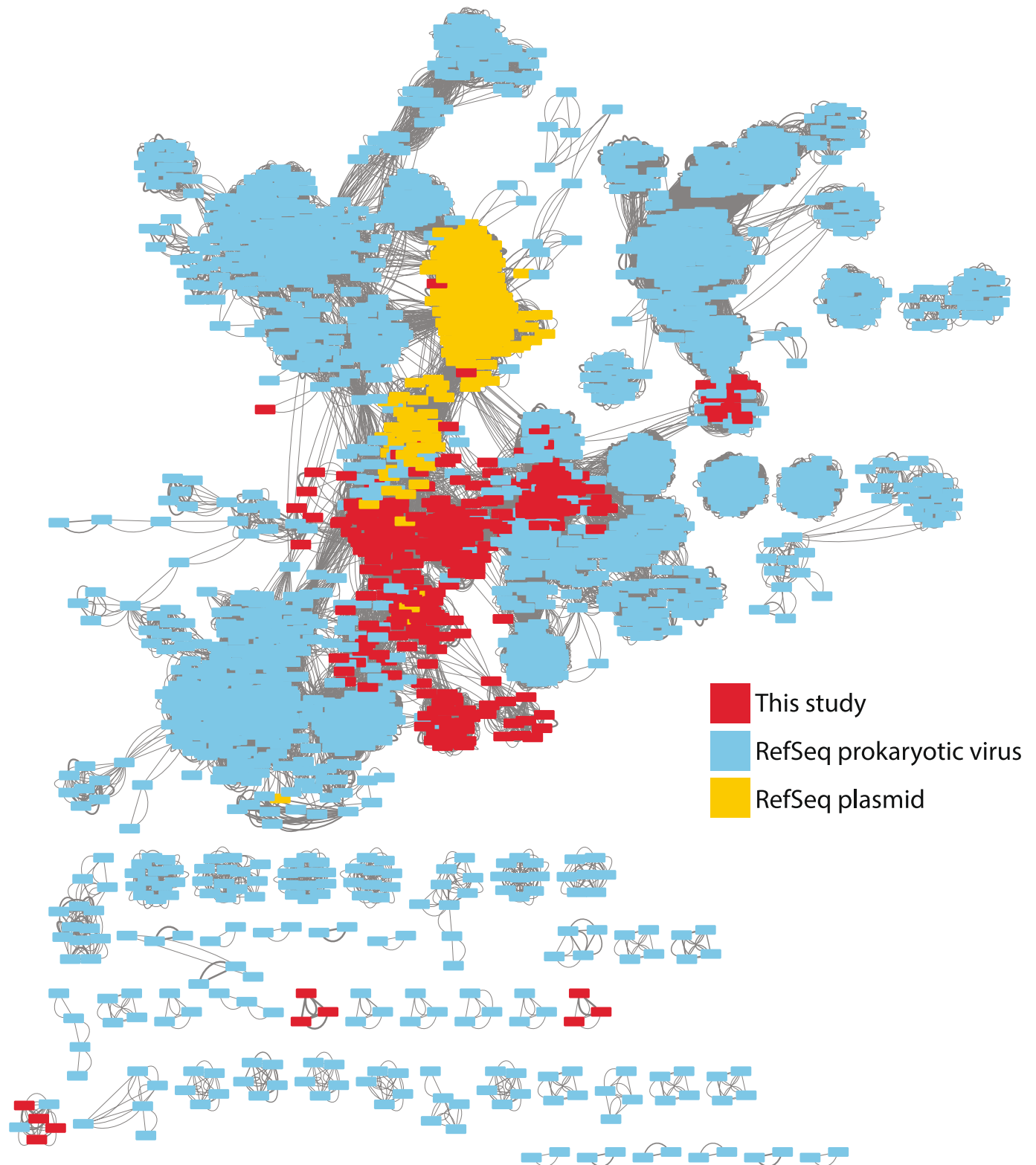
Comparisons of gene predictions for a region with genes of clearly predicted function in M05_PHAGE_COMPLETE_32_3. **a**, The standard (code 11) genetic code. **b**, Both TAG and TAA repurposed (code 6). **c**, TAG repurposed (code 16).

Overall, analysis of well-annotated genes supported code 16 as the best choice (TAG to X, as X could not be clearly resolved on the basis of sequence alignments with related proteins).



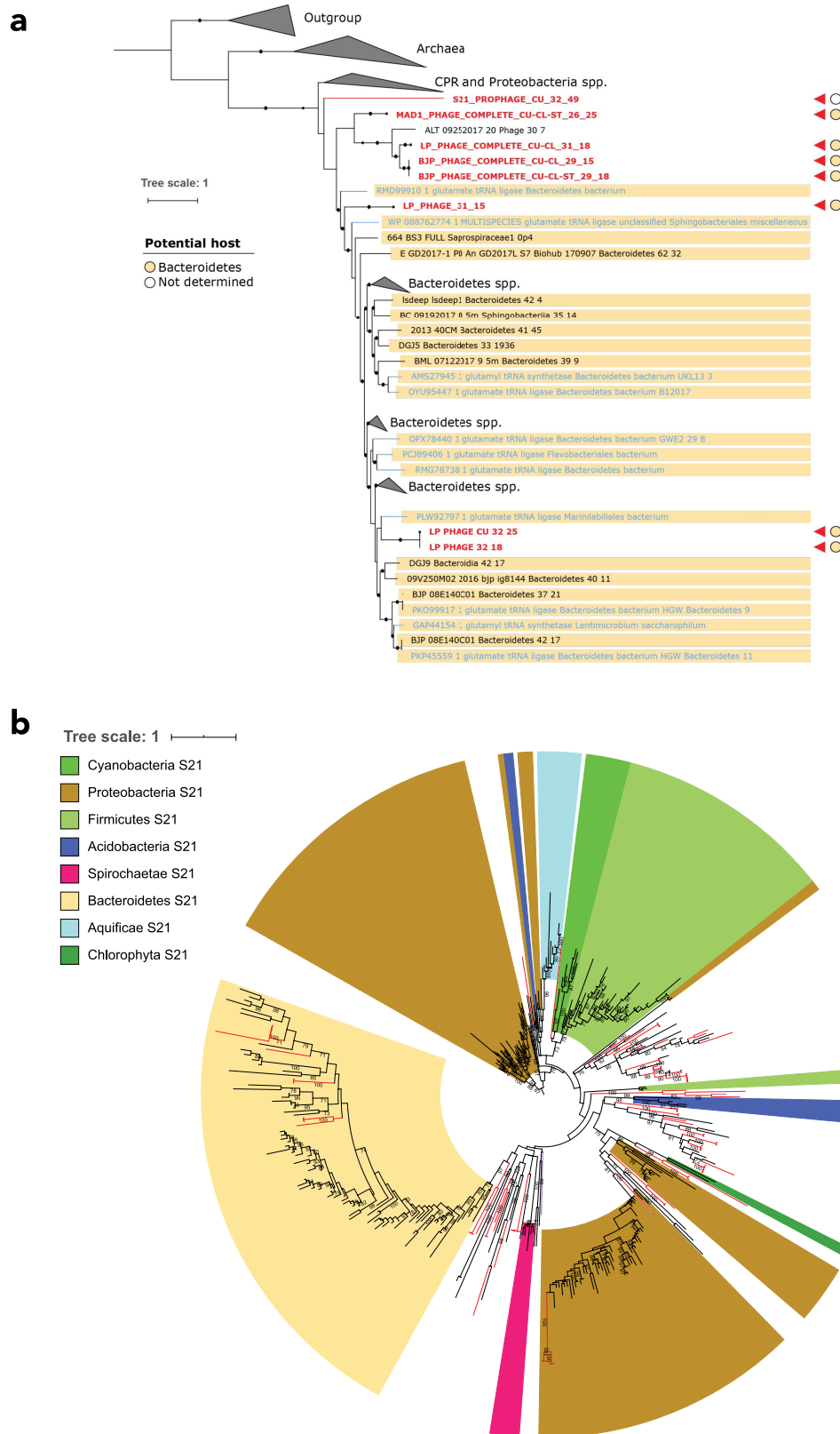
Extended Data Fig. 5 | Phylogenetic and protein-cluster relationships between phages. a, The phylogenetic tree of phages based on the MCPs. The outer ring shows genome length; bars in red indicate genomes reconstructed and reported in this study and bars in blue indicate database genomes. The next ring indicates the environment of origin. The inner ring indicates the phylum of the host (black indicates unknown). Superimposed colours indicate named clades that consist of huge phages that were identified in the terminase tree. Colours are as in Fig. 2. **b,** Hierarchical clustering dendrogram of phage

genomes based on the Jaccard distance between the presence or absence profiles of protein families, performed using an average linkage method. The outermost ring shows phage genome length, the next ring shows the environment of origin, then predicted phylum affiliation of bacterial hosts. Superimposed colours indicate named clades that consist of huge phages that were identified in the terminase tree. Colours are as in Fig. 2. The clustering supports the phylogenetic analyses shown in **a** and Fig. 2.



Extended Data Fig. 6 | Protein-clustering network for phages and plasmids. Network analysis using vContact2 and Cytoscape⁸⁶ based on the number of shared protein clusters between the genomes in this study, RefSeq prokaryotic virus genomes and 400 randomly sampled plasmid sequences from RefSeq.

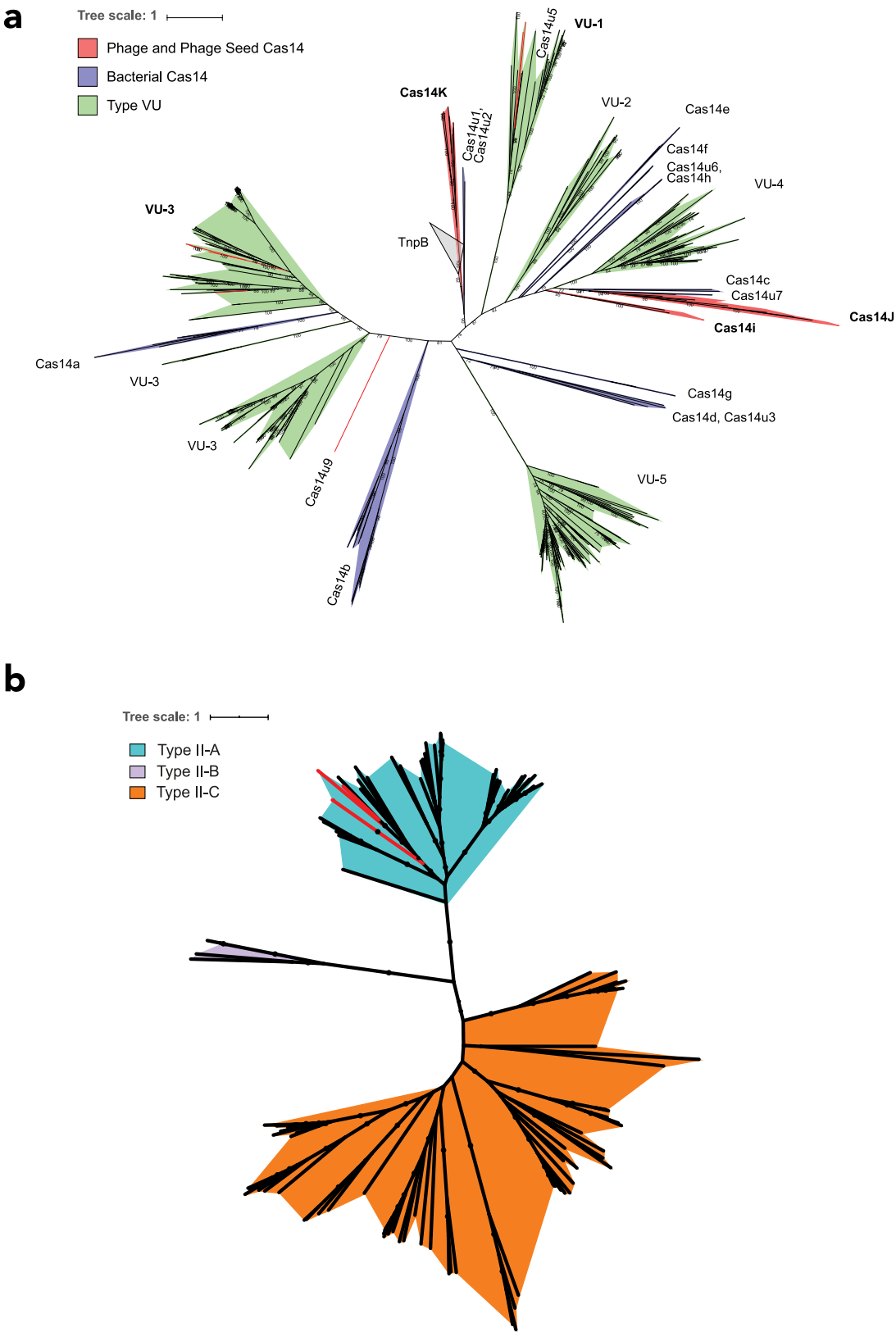
Each node represents a genome and each edge is the hypergeometric similarity (>30) between genomes based on shared protein clusters. This analysis was used to help to distinguish between the classification of genomes as phage, plasmid or unknown.



Extended Data Fig. 7 | Phylogenetic analysis of tRNA synthetase.

a, Aminoacyl tRNA synthetases were detected in many huge phages reported in this study (Supplementary Table 6). The phylogenetic subtree for glutamate-tRNA synthetase sequences from phages (red text and small triangles) that place within or close to sequences from Bacteroidetes hosts is shown as an example. Bacterial sequences from public databases are indicated by black text

and those from metagenomes from which huge phage genomes were reconstructed are indicated by blue text. Coloured circles indicate the predicted phylum of the bacterial host for each phage. **b**, Phylogenetic tree of phage-encoded ribosomal protein S21 and the top RefSeq hits for each protein, constructed using IQTREE. Sequences from this study are indicated by red branches.

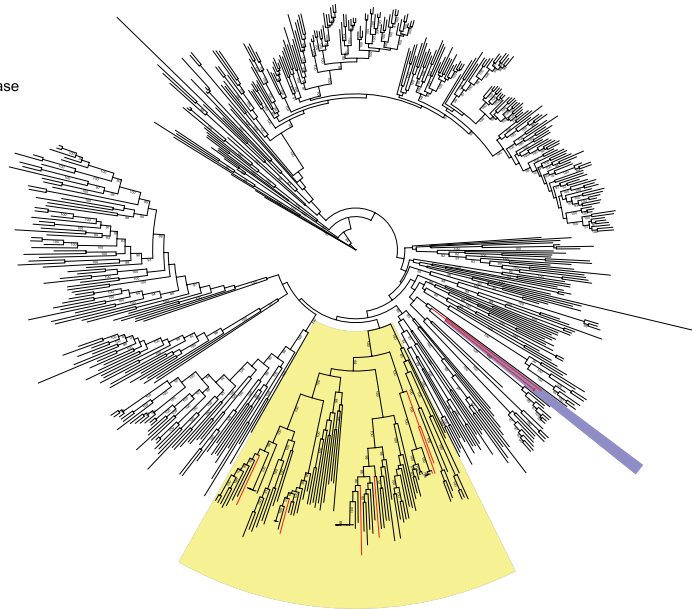


Extended Data Fig. 8 | Phylogenetic trees of Cas14, CRISPR–Cas type V-U and Cas9. a, Phylogenetic tree for Cas14 and type V-U. **b,** Phylogenetic tree for Cas9. Sequences from this study are indicated by red branches.

a**b**

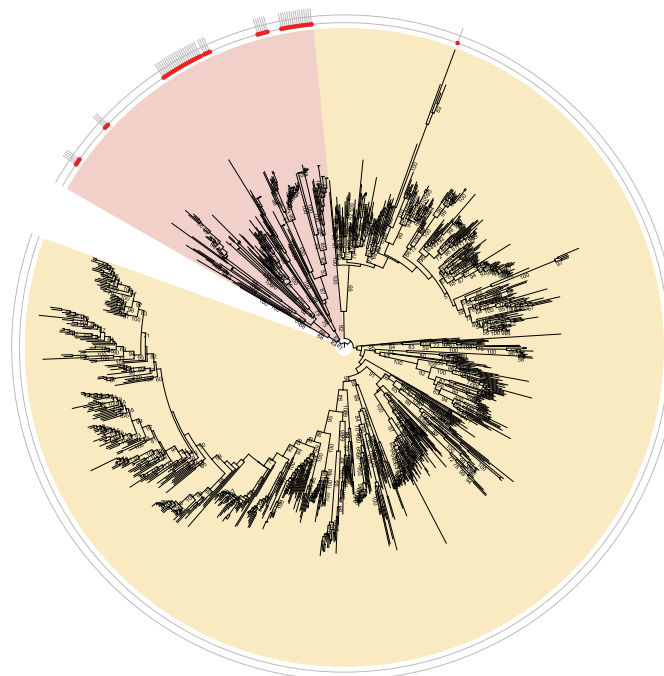
Tree scale: 1 ———

■ Cas3
■ Unidentified Helicase

**c**

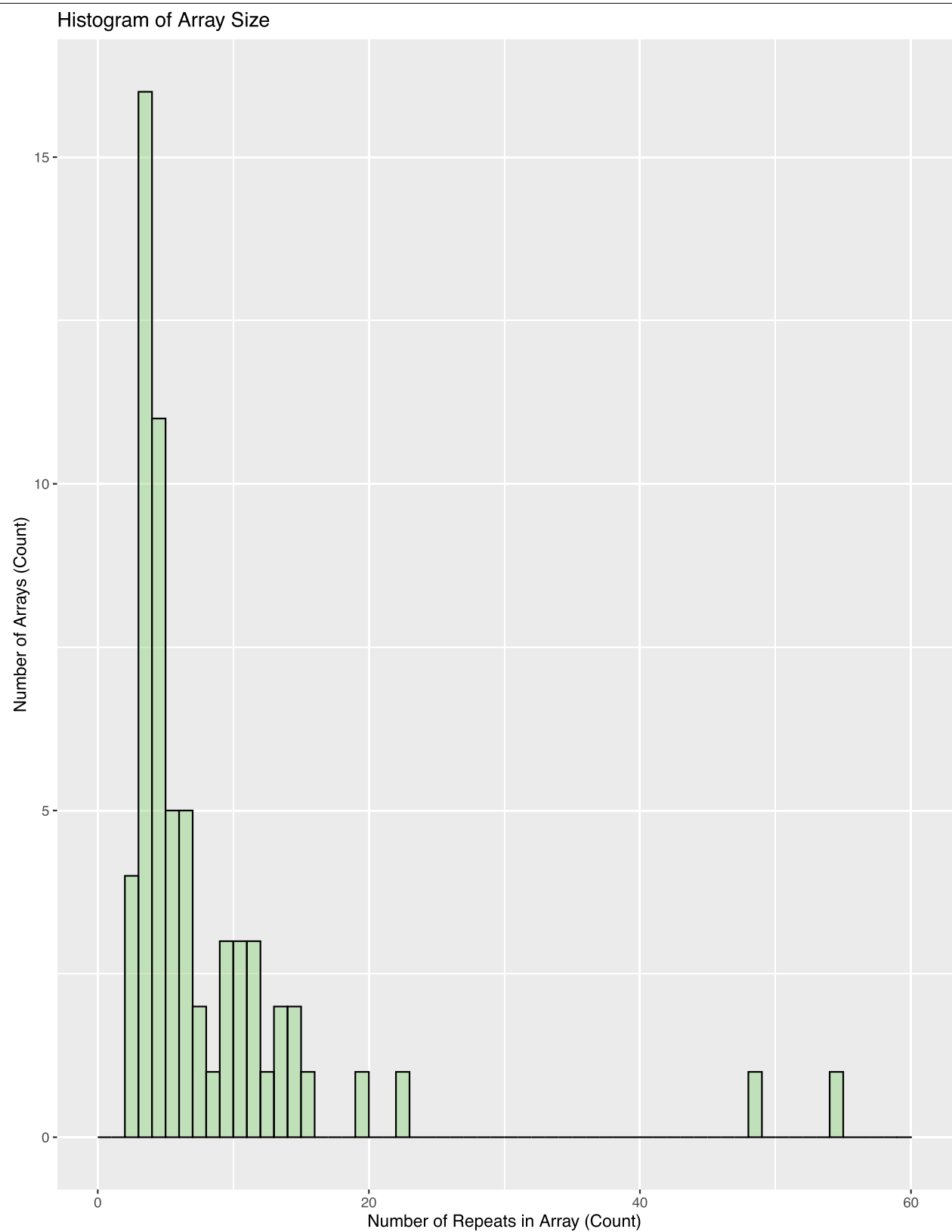
Tree scale: 1 ———

■ Cas4-like proteins
■ Cas4



Extended Data Fig. 9 | Variant type I CRISPR–Cas system and Cas4-like proteins found in the genomes of huge phages. **a**, Locus architecture for type-I variant CRISPR phages. An interesting type-I system identified in huge phages lacks Cas6 but has Cas5, which is most similar to the Cas5d protein from type I-C, in which Cas5d acts as the pre-crRNA endonuclease (a role commonly reserved for Cas6). The proposed active site residues of Cas5d are to some extent different in the Cas5 of this system, although this may still confer

processing activity, as this change is also observed in other Cas6 homologues. **b**, Phylogenetic tree of superfamily 1–6 helicases, including Cas3 and the unidentified helicase in the type I-C variant system. Sequences from this study are indicated by red branches. **c**, Phylogenetic tree of Cas4, Cas4-like proteins from the phage and plasmid genomes reported here, and the top 50 RefSeq hits to the Cas4-like proteins. Cas4-like genes from this study are denoted by red circles.



Extended Data Fig. 10 | Distribution of phage- and plasmid-encoded CRISPR array sizes. The indicated count is the number of recovered repeats.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | | |
|-------------------------------------|--|
| n/a | Confirmed |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Geneious v9.1.8 (Licensed, paid version used in this study, free versions available)
BBmap v37.5
IDBA_UD v1.1.1
Bowtie2 v2.3.4.1
MEGAHIT v1.1.3
SPAdes v3.11.1

Data analysis

vContact2
Vmatch v2.3
Prodigal v2.6.3
tRNAscan-SE v2.0
ARAGORN v1.2.38
MMseqs2 Version: 9f493f538d28b1412a2d124614e9d6ee27a55f45
HHSuite v3.0.3
CD-HIT v4.6.8
SignalP v4.1
DAMA v1.0
PSORT v3.0
TMHMM v2.0
MAFFT v7.407
DIAMOND v0.9.24
RAXML v8.0.26
IQTREE v1.6.6
HMMER v3.1b2
GLIMMER3 v1.5

MinCED v0.2.0
https://github.com/rohansachdeva/assembly_repeats v0

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Accession numbers for each genome and reads are provided in the data availability statement and in Table S1. Genbank files for each genome are also provided as supplementary data.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The sample size was chosen to provide a high breadth of ecosystem coverage for the recovery of huge phage genomes. Accordingly, there were no statistical methods to determine sample sizes. Data were compiled from multiple sources where virus-like genomes could be found. Data were compiled from multiple sources wherever virus-like genomes could be found. Samples from each sampling site is listed in Table 1 and Table S1
Data exclusions	IMG/VR phage data was excluded because of the prevalence of artifactual concatenated viral sequence assemblies. This was pre-established based on criteria in Devoto et al. 2019.
Replication	Near identical phage genomes were recovered from multiple independent samples, verifying sequence assembly and genome reconstruction. Host identification was verified by a combination of CRISPR targeting, phylogenetic analysis of ribosomal proteins, and phylum-level taxonomic profiles. Annotations were verified across multiple databases.
Randomization	Randomization is not applicable to this because the reported study is a survey of huge phage genomes across many ecosystems and not dependent on experimental outcomes.
Blinding	Blinding was not performed because it was not applicable to this study. This study was a survey of huge phage genomes across global populations, and was not dependent on trial outcomes

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- | | |
|-------------------------------------|--|
| n/a | Involved in the study |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |

Methods

- | | |
|-------------------------------------|---|
| n/a | Involved in the study |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

Giant virus diversity and host interactions through global metagenomics

<https://doi.org/10.1038/s41586-020-1957-x>

Received: 4 June 2019

Accepted: 9 January 2020

Published online: 22 January 2020

Open access

 Check for updates

Frederik Schulz^{1✉}, Simon Roux¹, David Paez-Espino¹, Sean Jungbluth¹, David A. Walsh², Vincent J. Denef³, Katherine D. McMahon^{4,5}, Konstantinos T. Konstantinidis⁶, Emiley A. Eloë-Fadrosch¹, Nikos C. Kyrpides¹ & Tanja Woyke^{1✉}

Our current knowledge about nucleocytoplasmic large DNA viruses (NCLDV) is largely derived from viral isolates that are co-cultivated with protists and algae. Here we reconstructed 2,074 NCLDV genomes from sampling sites across the globe by building on the rapidly increasing amount of publicly available metagenome data. This led to an 11-fold increase in phylogenetic diversity and a parallel 10-fold expansion in functional diversity. Analysis of 58,023 major capsid proteins from large and giant viruses using metagenomic data revealed the global distribution patterns and cosmopolitan nature of these viruses. The discovered viral genomes encoded a wide range of proteins with putative roles in photosynthesis and diverse substrate transport processes, indicating that host reprogramming is probably a common strategy in the NCLDVs. Furthermore, inferences of horizontal gene transfer connected viral lineages to diverse eukaryotic hosts. We anticipate that the global diversity of NCLDVs that we describe here will establish giant viruses—which are associated with most major eukaryotic lineages—as important players in ecosystems across Earth's biomes.

Large and giant viruses of the NCLDV supergroup have complex genomes with sizes of up to several megabases, and virions that are a similar size to, or even larger than, small cellular organisms^{1–3}. These viruses infect a wide range of eukaryotes from protists to animals⁴. Marker gene surveys have shown that NCLDVs are not only extremely abundant and diverse in oceans^{5–7}, but can also frequently be found in freshwater⁸ and soil⁹. However, the discovery of large and giant viruses has mainly been driven by their co-cultivation with amoebae or isolation together with their native hosts^{1,4,8}. Only recently, metagenomic and single-cell genomic studies have facilitated the discovery of several new NCLDV members and showed that cultivation-independent methods are applicable to these viruses just as they are to uncultivated Bacteria and Archaea^{9–14}.

Here, we have used a multistep metagenome data-mining, binning and iterative-filtering pipeline (Extended Data Figs. 1, 2 and Supplementary Text 1), which led to the recovery of genomes representing 2,074 putative NCLDV populations from 8,535 publicly available metagenomes in the Integrated Microbial Genomes and Microbiomes (IMG/M) database¹⁵. The assembly size, GC content, coding density and copy number of nucleocytoplasmic virus orthologous genes (NCVOGs)¹⁶ were comparable to previously described NCLDV genomes, supporting the classification of these genomes as giant virus metagenome-assembled genomes (GVMAGs) (Extended Data Figs. 3, 4 and Supplementary Tables 1–3). Using an approach that relied on conserved NCVOGs, we estimated genome completeness and contamination, which led to the classification of 773 high-quality, 989 medium-quality and 312 low-quality GVMAGs (Extended

Data Figs. 1, 4 and Supplementary Tables 1, 4), in line with the MIUViG recommendations¹⁷.

Augmenting the existing NCLDV phylogenetic framework with the GVMAGs substantially increased the diversity of this proposed viral order (Fig. 1a and Supplementary Data 1). The resulting phylogenetic tree expanded from 205 to 2,279 viral genomes, which can now be divided into 100 potentially genus- or subfamily-level monophyletic clades spanning 10 provisional superclades, compared with the previously recognized 20 genera². This translates into an 11-fold increase in phylogenetic diversity of the NCLDVs. Notably, the addition of the novel viral genomes did not change the basic topology of the NCLDV tree but rather altered the contribution of existing groups, the *Mimiviridae* in particular, to the total viral diversity. Furthermore, the presence of conserved NCVOGs in lineage-specific patterns strengthens the hypothesis of a common evolutionary origin of this viral group². Novel groups of viruses with no isolate representatives appeared within the existing taxonomic framework (that is, metagenomic giant virus lineages (MGVLs)). The greatest number of GVMAGs could be attributed to MGVL57 ($n = 205$), the Yellowstone Lake mimiviruses (YLMVs; $n = 119$) and MGVL42 ($n = 84$). In addition, several established viral lineages were considerably extended, such as the prasinoviruses ($n = 77$), iridoviruses ($n = 59$), cafeteriaviruses ($n = 43$), phaeocystisviruses ($n = 37$), klosneuviruses ($n = 36$), tetraselmisviruses ($n = 34$) and raphidoviruses ($n = 26$), some of which previously consisted of single isolates. In total, the GVMAGs increased the 123,000 previously known NCLDV proteins that clustered in 47,700 protein families to more than 924,000 proteins in 508,000 protein families (Extended Data Fig. 5a). Pfam-A protein

¹DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. ²Groupe de recherche interuniversitaire en limnologie, Department of Biology, Concordia University, Montréal, Québec, Canada. ³Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA. ⁴Department of Bacteriology, University of Wisconsin-Madison, Madison, WI, USA. ⁵Department of Civil and Environmental Engineering, University of Wisconsin-Madison, Madison, WI, USA. ⁶School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, USA. ✉e-mail: fschulz@lbl.gov; twoyke@lbl.gov

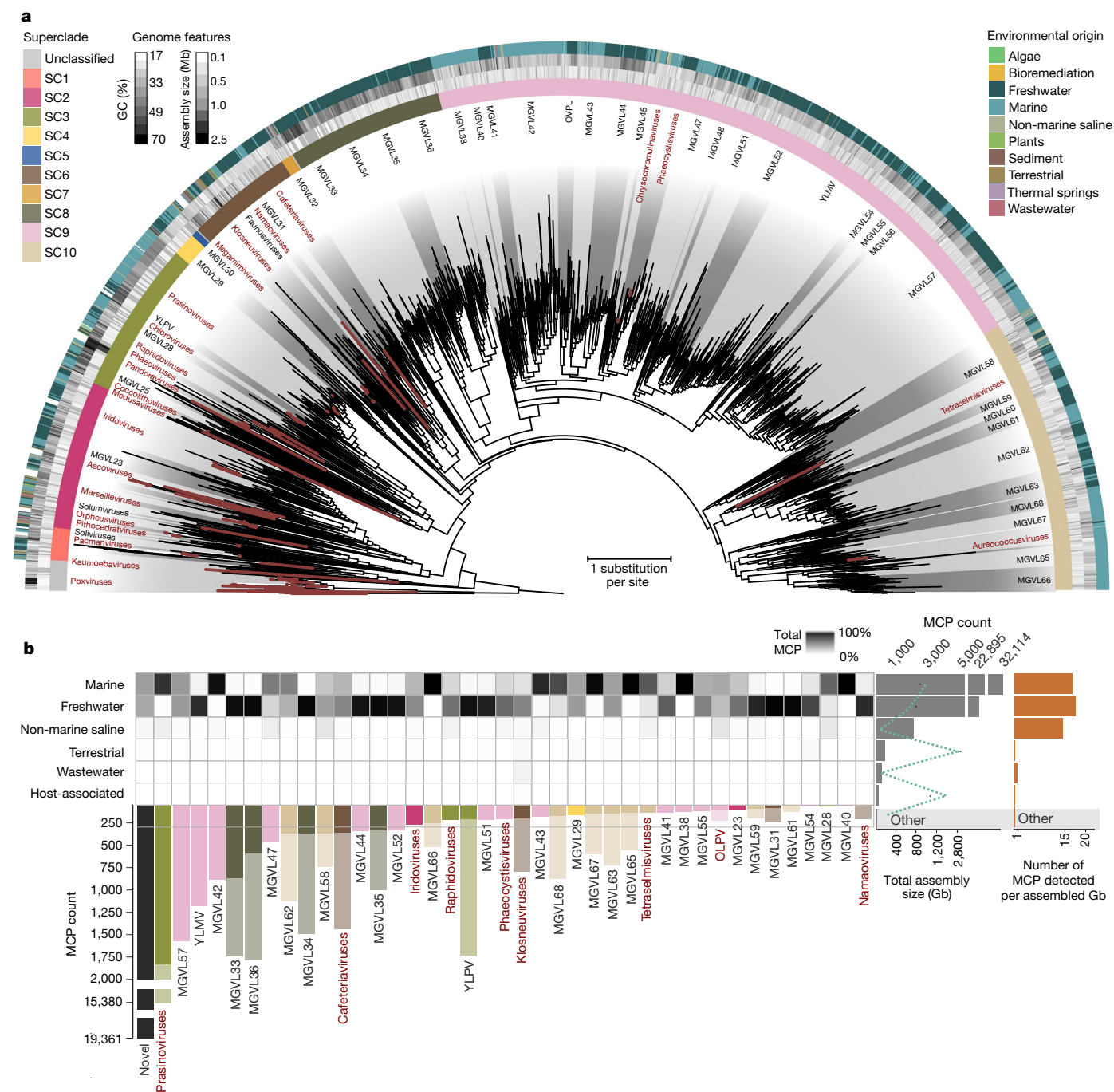


Fig. 1 | Metagenomic expansion of the NCLDV diversity. a, Maximum-likelihood phylogenetic tree of the NCLDV inferred from a concatenated protein alignment of five core NCVOGs¹⁶. Branches in dark red represent published genomes and branches in black represent GVMAGs generated in this study. Shades of grey indicate boundaries of genus- and subfamily-level clades; previously described lineages are labelled. Tree annotations from inside to the outside: (1) superclade (SC), (2) GC content, (3) assembly size and (4) environmental origin. **b**, Distribution of NCLDV lineages across different habitats. The bars adjacent to the heat map show the total number of detected

MCPs per habitat (facing to the right) and per lineage (facing downwards) as total count (total bar length) and corrected count on the basis of the average copy number of MCPs in the respective lineage (darker shaded bar length). The plot includes only lineages for which at least 100 MCPs could be detected. NCLDV lineages with available virus isolates are indicated in red. The turquoise dashed line indicates the total size of the metagenome assemblies that were screened in this analysis. Bars on the far right indicate, for each environment, the number of detected MCPs per assembled gigabase (Gb).

domains could be assigned to less than one third (31%) of these proteins (Extended Data Fig. 5b). The potentially most-versatile viral lineage on the basis of known gene functions were the klosneuviruses, for which more than 1,200 different protein domains could be detected (Extended Data Fig. 5b). MGVL57, MGVL58, YLMVs and klosneuviruses were the most-diverse lineages on the basis of their overall gene content, as

indicated by a low number of shared protein families compared with the total number of protein families (Extended Data Fig. 5c). MGVL27, medusaviruses, sylvanviruses and MGVL24 represented the viral lineages with the highest genome novelty; for these lineages, on average, less than 15% of proteins showed similarity to known NCLDV proteins (Extended Data Fig. 6). Notably, clades that had been predominantly

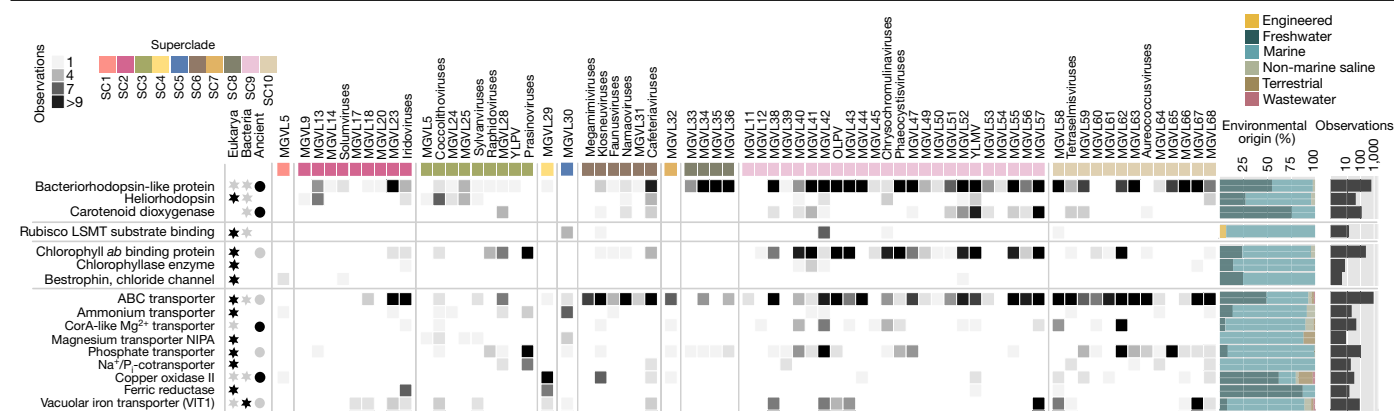


Fig. 2 | NCLDV coding potential and proteins that are probably involved in metabolic host reprogramming. Copy numbers of selected Pfam domains with potential roles as light-driven proton pumps, in carbon fixation, in photosynthesis and in diverse substrate transport processes. Filled stars and circles specify observed modes of transmission of the respective Pfam-domain-containing proteins. Stars represent recent HGTs from either

eukaryotes or bacteria; circles indicate vertical transmission after ancient HGT or gene birth in the NCLDV; a darker colour indicates the predominantly observed mode of transmission (five or more events). The stacked bars on the right side of the heat map show, for each observed protein domain, the proportional distribution across different habitat types. Bars on the far right indicate the total number of observations for each protein domain.

sampled in the past with several viral isolate genomes sequenced, such as marseilleviruses, poxviruses, pandoraviruses and faustoviruses, were nearly absent in the environmental microbiome data. This finding indicates that these viruses or their hosts have comparably low abundances in the samples analysed our dataset. It also suggests that there is a skew in the isolation and co-cultivation efforts of giant viruses using selected non-native hosts in laboratory setups^{18–20}. Large-scale, cultivation-independent genome-resolved metagenomics alleviates such bias and provides a more-global snapshot of diversity and the spatial distribution of NCLDVs in their natural habitats.

To further deepen our understanding of the environmental distribution patterns of the NCLDVs, we performed a survey of the major capsid protein (MCP) across all public metagenomic datasets. We identified more than 58,000 copies of this protein, of which 67% could be assigned to viral lineages (Fig. 1b). Among the most-commonly found lineages were prasinoviruses, MGVL57 and YLMV with more than 1,000 occurrences each. At the same time, only a few MCPs (less than 100) were detected in viruses that have repeatedly been isolated in co-cultivation with amoebae, such as megamimiviruses, marseilleviruses and faustoviruses^{18–20}. In our environmental survey, MCPs were predominantly found in marine (around 55%) and freshwater (about 40%) and—to a much lesser extent—in terrestrial (less than 1%) environments. Some NCLDV lineages occurred solely in either freshwater (YLMV, MGVL33 and MGVL36) or marine (prasinoviruses, MGVL42 and MGVL66) systems, whereas members of other lineages were found in both—or in an even-wider range of—environments (such as klosneuviruses, which were found in freshwater, marine, non-marine saline, terrestrial, wastewater and host-associated ecosystems). Large and giant viruses could also be detected in hydrothermal vents and thermal springs; however, comparably few MCPs were present in these habitats (Fig. 1b). Projecting the distribution of NCLDVs onto a global scale makes their ubiquitous nature apparent (Extended Data Fig. 7). These viruses can be found almost anywhere with many different lineages often co-occurring in close proximity to each other, suggesting that their discovery is chiefly limited by sampling effort.

Considering the ubiquitous prevalence of large and giant viruses, we aimed to investigate the potential influences that these viruses have on their hosts. The detrimental effect of viral infections on their eukaryotic hosts are well-known¹; however, a few recent studies have shown that NCLDVs might also complement the metabolism of their host, for example, by encoding transporters that take up nutrients, such as nitrogen, or fermentation genes^{21,22}. Expanding these initial findings,

our data showed that diverse lineages across all NCLDV superclades encoded enzymes with potential roles in photosynthesis, diverse substrate transport processes, light-driven proton pumps and retinal pigments (Fig. 2). Maps of the presence, absence and prevalence of these genes revealed lineage- and environment-specific patterns. Most-commonly observed across a wide-range of habitats were ABC transporters, chlorophyll *ab*-binding proteins and bacteriorhodopsin-like proteins (Fig. 2, Supplementary Note 2 and Supplementary Table 5). Transporters for ammonium, magnesium and phosphate, which are likely to be of importance for hosts in oligotrophic environments such as the surface ocean, were predominantly found in marine viruses. Enzymes such as ferric reductases and multicopper oxidases—which facilitate the uptake of iron^{23,24}, an essential trace element that is often growth-limiting, especially in photosynthetic organisms²⁵—were encoded in GVMAGs sampled across different habitats. This wealth of virus-encoded genes with roles in energy generation and nutrient acquisition has far-reaching implications for ecosystem dynamics. Metabolic reprogramming refers to a common phenomenon in which bacterial viruses obtain genes from their hosts and maintain them to support host metabolism²⁶. Our results illustrate that in a similar manner, NCLDV-mediated host reprogramming is probably an important strategy to increase viral fecundity and at the same time render a short-term competitive advantage of infected eukaryotic host cells, especially under nutrient-limited conditions.

In agreement with previous studies^{27–30}, many of the identified viral genes with predicted effects on host cell processes were probably acquired from their hosts through horizontal gene transfer (HGT) (Fig. 2 and Extended Data Fig. 8). Other genes were present across different viral lineages and superclades, suggesting ancient transfer followed by vertical inheritance during the course of NCLDV evolution or the origin of the respective gene in a common ancestor of this group of viruses. A notable example is the group of rhodopsin-like domain-containing proteins, which we found in 555 of the GVMAGs. Type-1 rhodopsins in algae-infecting phycodnaviruses and in viruses of heterotrophic choanoflagellates have been reported in previous studies and comprise viral rhodopsin groups I and II^{30,31,32}. However, in light of our extended sampling of NCLDV genomes, it becomes evident that NCLDVs encoded more-diverse rhodopsins than described (Extended Data Fig. 8), which comprise approximately one quarter of the total known diversity of rhodopsins and include proteins from all publicly available metagenomes (Extended Data Fig. 9). Notably, the phylogeny

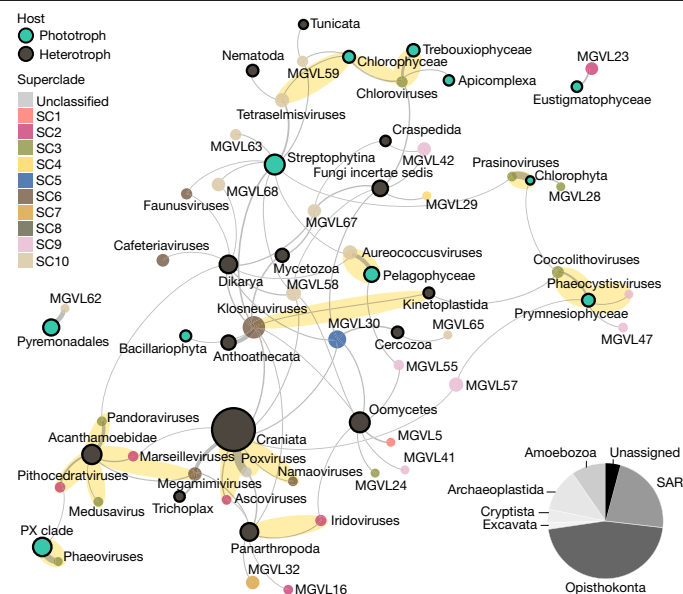


Fig. 3 | HGT between NCLDV and their putative eukaryotic hosts. Undirected HGT network with nodes that represent previously described viral lineages and MGVLs, coloured on the basis of NCLDV superclade affiliation, with names above the node and their putative hosts (highlighted in black with names below the node, coloured on the basis of lifestyle); edges are weighted on the basis of the number of detected transfers. Connections comprising at least four transfers are shown. Experimentally verified virus–host associations are highlighted in yellow with names in bold. The proportion of HGT candidates assigned to hosts from different major eukaryotic lineages is shown as a pie chart.

of the viral rhodopsins from all NCLDV superclades exhibits a strongly supported monophyletic signal, which implies that this gene might represent an ancestral trait of the NCLDV that was subsequently lost in some lineages. In addition to viral rhodopsin group I and II, additional NCLDV rhodopsins branch closely to their cellular counterparts and have probably been acquired by HGT from different hosts (Extended Data Fig. 8). In a similar manner, putative NCLDV heliorhodopsins were found intertwined with their homologues in the algae *Chrysochromulina* and *Micromonas* (Extended Data Fig. 8). In addition to the rhodopsins, our dataset contained 119 GVMAgs that encoded carotenoid oxygenases, which potentially modulate light-harvesting capacity or synthesize bioactive compounds³³. It is conceivable that some of the NCLDV rhodopsins function in conjunction with the carotenoid oxygenases and have important roles in modulating host-cell processes; for example, by acting as light-driven proton pumps, as photoreceptors in host phototactic motility or as photoprotectants^{10,34,35}—each of these functions lead to metabolic advantages of infected populations.

Uptake of host genes is a common mechanism in the evolution of NCLDVs^{2,11,30,36}. Using HGT analyses, we assigned putative hosts to different NCLDV lineages. Analysis of 2,040 genes that have probably undergone HGT provided linkage information for 50 viral lineages to 32 groups of putative eukaryotic hosts (Fig. 3 and Supplementary Table 6). Notably, 17 out of 23 viral lineages that contained genomes from isolated viruses could be connected through HGT to their experimentally verified native hosts, such as most algae-infecting viruses and metazoa-infecting ascoviruses, namaoviruses and poxviruses, as well as connecting klosneuviruses to Kinetoplastida^{37,38}. Our analysis further confirmed *Acanthamoeba* as a host of pandoraviruses, pithocedratviruses, medusaviruses, marseilleviruses and megamimiviruses. Notably, megamimiviruses, which have exclusively been obtained through co-cultivation with amoebae, showed not only HGT with this host but were linked even more strongly to multicellular animals. The best-connected NCLDV lineage was the klosneuviruses, a viral subfamily

mainly known from metagenomic studies^{9,11,12,39}. Our HGT network revealed that klosneuviruses have a diverse putative host range of mainly heterotrophs, including Anthoathecata—to which it showed the strongest connection—as well as fungi and arthropods, and different protists, including slime moulds. By contrast, Oomycetes, Dikarya, fungi incertae sedis and Streptophytina emerged as putative hosts for the greatest number of different NCLDV lineages, despite the lack of isolation of NCLDVs from any of these organisms. With predicted hosts in Opisthokonta, Amoebozoa, Excavata, Archaeplastida, Cryptista and the Stramenopila, Alveolata, Rhizaria (SAR) supergroup, our results suggest that members of the NCLDV might be able to infect most major eukaryotic lineages⁴⁰ (Fig. 3). This is consistent with previous reports based on eukaryotic genome data²⁷ and experimental data showing that large and giant viruses infect marine arrow worms⁴¹, epithelial cells in fish gills³⁸ and potentially also corals and sponges⁴². Of note, our analysis did not reveal linkage to human hosts. We expect that with improved sampling of host genomes—particularly genomes of underexplored protists and algae—host linkage through HGT will yield an even more comprehensive picture of the host range and evolutionary histories of NCLDVs.

Overall, we leveraged the availability of metagenomic data generated by the global sampling efforts of a community of scientists to expand our insights into the diversity, host metabolic complementation and putative host range of large and giant viruses. NCLDV infections probably occur in all major eukaryotic lineages, with repercussions for many of Earth's major biogeochemical processes. Our data and findings represent a solid foundation and expansive resource for future giant-virus research efforts to deepen our understanding of the evolutionary and ecological bearings of these viral giants.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-1957-x>.

1. Abergel, C., Legendre, M. & Claverie, J.-M. The rapidly expanding universe of giant viruses: mimivirus, pandoravirus, pithovirus and mollivirus. *FEMS Microbiol. Rev.* **39**, 779–796 (2015).
2. Koonin, E. V. & Yutin, N. Evolution of the large nucleocytoplasmic DNA viruses of eukaryotes and convergent origins of viral gigantism. *Adv. Virus Res.* **103**, 167–202 (2019).
3. Abrahão, J. et al. Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. *Nat. Commun.* **9**, 749 (2018).
4. Fischer, M. G. Giant viruses come of age. *Curr. Opin. Microbiol.* **31**, 50–57 (2016).
5. Mihara, T. et al. Taxon richness of “Megaviridae” exceeds those of Bacteria and Archaea in the ocean. *Microbes Environ.* **33**, 162–171 (2018).
6. Hingamp, P. et al. Exploring nucleocytoplasmic large DNA viruses in Tara Oceans microbial metagenomes. *ISME J.* **7**, 1678–1695 (2013).
7. Monier, A., Claverie, J.-M. & Ogata, H. Taxonomic distribution of large DNA viruses in the sea. *Genome Biol.* **9**, R106 (2008).
8. Wilson, W. H., Van Etten, J. L. & Allen, M. J. The Phycodnaviridae: the story of how tiny giants rule the world. *Curr. Top. Microbiol. Immunol.* **328**, 1–42 (2009).
9. Schulz, F. et al. Hidden diversity of soil giant viruses. *Nat. Commun.* **9**, 4881 (2018).
10. Needham, D. M. et al. A distinct lineage of giant viruses brings a rhodopsin photosystem to unicellular marine predators. *Proc. Natl Acad. Sci. USA* **116**, 20574–20583 (2019).
11. Schulz, F. et al. Giant viruses with an expanded complement of translation system components. *Science* **356**, 82–85 (2017).
12. Bäckström, D. et al. Virus genomes from deep sea sediments expand the ocean megavirome and support independent origins of viral gigantism. *mBio* **10**, e02497-18 (2019).
13. Andreani, J., Verneau, J., Raoult, D., Levasseur, A. & La Scola, B. Deciphering viral presences: two novel partial giant viruses detected in marine metagenome and in a mine drainage metagenome. *Virology* **15**, 66 (2018).
14. Wilson, W. H. et al. Genomic exploration of individual giant ocean viruses. *ISME J.* **11**, 1736–1745 (2017).
15. Chen, I. A. et al. IMG/M v5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res.* **47**, D666–D677 (2019).
16. Yutin, N., Wolf, Y. I., Raoult, D. & Koonin, E. V. Eukaryotic large nucleocytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. *Virology* **6**, 223 (2009).
17. Roux, S. et al. Minimum Information about an Uncultivated Virus Genome (MIUViG). *Nat. Biotechnol.* **37**, 29–37 (2019).

18. Aherfi, S., Colson, P., La Scola, B. & Raoult, D. Giant viruses of amoebas: an update. *Front. Microbiol.* **7**, 349 (2016).
19. Boughalmi, M. et al. High-throughput isolation of giant viruses of the Mimiviridae and Marseilleviridae families in the Tunisian environment. *Environ. Microbiol.* **15**, 2000–2007 (2013).
20. Reteno, D. G. et al. Faustovirus, an asfarvirus-related new lineage of giant viruses infecting amoebae. *J. Virol.* **89**, 6585–6594 (2015).
21. Monier, A. et al. Host-derived viral transporter protein for nitrogen uptake in infected marine phytoplankton. *Proc. Natl Acad. Sci. USA* **114**, E7489–E7498 (2017).
22. Schvarcz, C. R. & Steward, G. F. A giant virus infecting green algae encodes key fermentation genes. *Virology* **518**, 423–433 (2018).
23. Saikia, S., Oliveira, D., Hu, G. & Kronstad, J. Role of ferric reductases in iron acquisition and virulence in the fungal pathogen *Cryptococcus neoformans*. *Infect. Immun.* **82**, 839–850 (2014).
24. Herbig, A., Bölling, C. & Buckhout, T. J. The involvement of a multicopper oxidase in iron uptake by the green algae *Chlamydomonas reinhardtii*. *Plant Physiol.* **130**, 2039–2048 (2002).
25. Morrissey, J. & Bowler, C. Iron utilization in marine cyanobacteria and eukaryotic algae. *Front. Microbiol.* **3**, 43 (2012).
26. Hurwitz, B. L., Hallam, S. J. & Sullivan, M. B. Metabolic reprogramming by viruses in the sunlit and dark ocean. *Genome Biol.* **14**, R123 (2013).
27. Gallot-Lavallée, L. & Blanc, G. A glimpse of nucleocytoplasmic large DNA virus biodiversity through the eukaryotic genomics window. *Viruses* **9**, 17 (2017).
28. Finke, J. F., Winget, D. M., Chan, A. M. & Suttle, C. A. Variation in the genetic repertoire of viruses infecting *Micromonas pusilla* reflects horizontal gene transfer and links to their environmental distribution. *Viruses* **9**, 116 (2017).
29. Maumus, F. & Blanc, G. Study of gene trafficking between *Acanthamoeba* and giant viruses suggests an undiscovered family of amoeba-infecting viruses. *Genome Biol. Evol.* **8**, 3351–3363 (2016).
30. Filée, J. & Chandler, M. Gene exchange and the origin of giant viruses. *Intervirology* **53**, 354–361 (2010).
31. Philosofo, A. & Béjà, O. Bacterial, archaeal and viral-like rhodopsins from the Red Sea. *Environ. Microbiol. Rep.* **5**, 475–482 (2013).
32. Yutin, N. & Koonin, E. V. Proteorhodopsin genes in giant viruses. *Biol. Direct* **7**, 34 (2012).
33. Ahrazem, O., Gómez-Gómez, L., Rodrigo, M. J., Avalos, J. & Limón, M. C. Carotenoid cleavage oxygenases from microbes and photosynthetic organisms: features and functions. *Int. J. Mol. Sci.* **17**, 1781 (2016).
34. Ernst, O. P. et al. Microbial and animal rhodopsins: structures, functions, and molecular mechanisms. *Chem. Rev.* **114**, 126–163 (2014).
35. Sineshchekov, O. A., Jung, K.-H. & Spudich, J. L. Two rhodopsins mediate phototaxis to low- and high-intensity light in *Chlamydomonas reinhardtii*. *Proc. Natl Acad. Sci. USA* **99**, 8689–8694 (2002).
36. Moreira, D. & Brochier-Armanet, C. Giant viruses, giant chimeras: the multiple evolutionary histories of Mimivirus genes. *BMC Evol. Biol.* **8**, 12 (2008).
37. Deeg, C. M., Chow, C. T. & Suttle, C. A. The kinetoplastid-infecting Bodo saltans virus (BsV), a window into the most abundant giant viruses in the sea. *eLife* **7**, e33014 (2018).
38. Clouthier, S., Anderson, E., Kurath, G. & Breyta, R. Molecular systematics of sturgeon nucleocytoplasmic large DNA viruses. *Mol. Phylogenet. Evol.* **128**, 26–37 (2018).
39. Stough, J. M. A. et al. Diversity of active viral infections within the *Sphagnum* microbiome. *Appl. Environ. Microbiol.* **84**, e01124-18 (2018).
40. Adl, S. M. et al. The revised classification of eukaryotes. *J. Eukaryot. Microbiol.* **59**, 429–514 (2012).
41. Shinn, G. L. & Bullard, B. L. Ultrastructure of Meelsvirus: a nuclear virus of arrow worms (phylum Chaetognatha) producing giant “tailed” virions. *PLoS ONE* **13**, e0203282 (2018).
42. Claverie, J.-M. et al. Mimivirus and Mimiviridae: giant viruses with an increasing number of potential hosts, including corals and sponges. *J. Invertebr. Pathol.* **101**, 172–180 (2009).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

Methods

Generation of models to detect NCLDV proteins

Initial hidden Markov models (HMMs) for the MCPs were built from a multiple sequence alignment of published NCLDV MCPs and subsequently updated on the basis of extracted metagenomic NCLDV MCP sequences. We screened around 537 million proteins encoded on about 45.1 million contigs with a length greater than 5 kb available in 8,535 public metagenomes in IMG/M⁴³ (June 2018) for contigs that encode the NCLDV MCP using a version of *hmmsearch* (v.3.1b2, <http://hmmer.org/>) that is optimized⁴⁴ for the supercomputer Cori, with a set of models for the NCLDV MCP (<https://bitbucket.org/berkeleylab/mtg-gv-exp/>) and an *E*-value cut-off of 1×10^{-10} . The 1,003,222 proteins found on the 77,701 contigs with hits for MCPs were then clustered with CD-hit⁴⁵ at a sequence similarity of 99% to remove nearly identical and identical proteins. This resulted in 524,161 clusters and singletons. The cluster representatives were used to infer protein families using orthofinder (v.2.27) with default settings and the *-diamond* flag^{46,47}. Multiple sequence alignments were built with mafft⁴⁸ (v.7.294b) for protein families that included at least 10 members and corresponding HMM models were obtained with *hmmbuild* (v.3.1b2, <http://hmmer.org/>). This led to a total of 7,182 HMMs that can detect NCLDV proteins that were then tested against all public genomes in IMG/M⁴³ (June 2018). Models that gave rise to hits above an *E*-value cut-off of 1×10^{-10} in more than 10 reference genomes were removed. The resulting 5,064 models were then used for targeted binning of NCLDV metagenome contigs.

Identification of NCLDV-specific genome features and design of an automatic classifier

A set of representative genomes of bacteria, archaea, eukaryotes and non-NCLDV viruses was gathered from the IMG/M database⁴³ (June 2018) and combined with NCLDV genomes assembled from metagenomes and protist genomes downloaded from NCBI GenBank to identify NCLDV-specific genome features. Genes were predicted for these genomes using Prodigal⁴⁹ (v.2.6.3; February, 2016) in both 'regular' mode (default parameters) and with the option '-n' activated, which forces a full motif scan. For genomes of less than 100 kb, the option '-p meta' was used to apply precalculated training files rather than training the gene predictor from the genome, as recommended by the tool documentation. Next, a set of different metrics was calculated for each genome on the basis of the genes predicted with a confidence of ≥ 90 and score of ≥ 50 . These included gene density (number of genes predicted on average per 10 kb of genome), coding density (number of bp predicted as part of a coding sequence per 10 kb of genome), spacer length (average length of the spacer between the predicted ribosomal binding site (RBS)), predicted start codon for genes in which a putative RBS was detected and RBS motif profile (the proportion of each type of RBS predicted in the genome, see below).

For the RBS motif profile, motifs were predicted using the full motif scan option of prodigal (see above). Notably, some of these motifs may not represent true RBSs, but are instead other conserved motifs (including transcription-related motifs) found upstream of start codons in these different genomes. These motifs were grouped into 11 categories as follows: (1) 'None' for cases in which prodigal did not predict a RBS; (2) 'SD_Canonical' for different variations of the canonical AGGAGG Shine–Dalgarno sequence (for example, AGGAG, AGxAG, GAGGA, as well as motifs identified by Prodigal as '3Base_5BMM' or '4Base_6BMM'); (3) 'SD_Bacteroidetes' for variations of the motif predicted typically from Bacteroidetes genomes (TA{2,5}T{0,1}: T followed by 2–5 As, and with sometimes a terminal T); (4) 'Other_GA' for motifs that include 'GA' patterns but that are different from the canonical Shine–Dalgarno sequence, for example, GAGGGA, typically identified in a few archaeal and bacterial genomes; (5) 'TATATA_3.6' for variations of the motif typically detected in NCLDV, that is, a motif of 3–6 bp with

alternating Ts and As (TAT, ATAT, TATA, TATAT, and so on); (6) 'OnlyA' for motifs exclusively composed of As not already included in a previous group, for example, AAAAA, most often found in Bacteroidetes; (7) 'OnlyT' for motifs exclusively composed of Ts not already included in a previous group, for example, TTTTT, found at a low frequency in some archaeal genomes; (8) 'DoubleA' for motifs with two consecutive As not already included in a previous group, for example, AAAAC, most often found in Bacteroidetes and bacteria from the candidate phyla radiation (CPR) group; (9) 'DoubleT' for motifs with two consecutive Ts not already included in a previous group, for example, TACTT, found at a low frequency in plants, Bacteroidetes and NCLDV; (10) 'NoA' for motifs without any As and not included in a previous group, for example, TCTCG, found in some archaeal genomes; and (11) 'Other' for motifs that did not fit into any of these categories.

Representative genomes were then grouped on the basis of the frequency of each motif type through hierarchical clustering (R function 'hclust'). This enabled the delineation of 12 genome groups on the basis of taxonomy (at the kingdom or domain ranks) and motif profile (Extended Data Fig. 2). Two types of random-forest classifiers were then built on the basis of the 14 features (11 motifs, gene density, coding density and average spacer length, see above): one for which the category to be predicted was binary (that is, 'Virus_NCLDV' versus 'Other') and one for which the category to be predicted was the set of genome groups based on predicted RBS motifs ('NCLDV (non-pandoraviruses)', 'animal and plants', 'protists & fungi', 'canonical bacteria and archaea', 'bacteroidetes-like', 'bacteria (CPR)', 'atypical bacteria', 'atypical archaea', 'plasmids' and 'other viruses', which include pandoraviruses). The 14 features were evaluated on the whole genomes, as well as on fragments of 20 kb and 10 kb selected randomly along the genomes. These random fragments were used to train a classifier on input sequences more comparable to metagenome assemblies, which most often represent short genome fragments of a few kb. For these fragments, Prodigal was run with the '-p meta' option and default parameters otherwise⁵⁰, that is, without a full motif scan, as these sequences are typically too short to identify de novo RBS motifs. Animal and plant genomes were not included in this analysis as these are highly unlikely to be assembled from metagenomes. All classifiers were built using R library randomforest and included 2,000 trees, with default parameters otherwise, and 10-fold cross-validation was performed to evaluate the classifier accuracy. The probability 'prob' of NCLDV origin was used as a prediction score to evaluate the classifiers and was then applied to metagenome assemblies. Because the input dataset is easily skewed towards bacterial and archaeal genomes, specificity and sensitivity were evaluated separately for each group of genome (Extended Data Fig. 2c). Statistical tests were performed in R using the package stats (Kolmogorov–Smirnov test)⁵¹ and effsize (Cohen's effect size)⁵².

MAGs from non-targeted binning of IMG genomes

Complementary to the targeted binning of NCLDV contigs, we performed genome binning of public metagenomes in IMG/M (assessed June 2018)¹⁵ with MetaBAT (v.0.32.4)⁵³ in the 'superspecific' mode, using read coverage information, if available in IMG, and a minimum contig length of 5 kb. Resulting MAGs were then checked for quality using CheckM (v.1.0.7)⁵⁴. Genome bins with completeness <50% were labelled as low quality according to the 'minimum information for a MAG' (MIMAG) standards⁵⁵.

Targeted binning of putative NCLDV metagenome contigs

The 5,064 NCLDV-specific models were used for *hmmsearch* (v.3.1b2, <http://hmmer.org/>) on the initial set of around 537 million proteins encoded on about 45 million contigs with a length greater than 5 kb with an *E*-value cut-off of 1×10^{-10} (Extended Data Fig. 1). In addition to the screening of the metagenomic contigs with NCLDV-specific models, we also used an automatic classifier using gene density and RBS motifs

Article

(see above). On the basis of the output of the automatic classifier, a score was assigned to each contig: a score of 2 if $\text{Ratio_TATATA_36} > 0.3$ or $\text{Pred_simple_NCLDV_score} > 0.3$ and the prediction result was 'Virus_NCLDV', a score of 1 if $\text{Ratio_TATATA_36} > 0.3$ or $\text{Pred_simple_NCLDV_score} > 0.1$ or the prediction result was 'Virus_NCLDV', otherwise a score of 0. On the basis of the cross-validation of the classifier, these parameters were chosen to maximize sensitivity while retaining enough specificity. The resulting set of around 1.2 million contigs with an RBS score of at least 1 and/or at least 20% of encoded genes (1 out of 5) with hits to the NCLDV models were subject to metagenomic binning as follows: for each metagenome, putative NCLDV contigs were extracted and binning performed with MetaBAT⁵⁶ (v.2) and contig read coverage information was used as input in case it was available in IMG⁴³. The targeted binning approach gave rise to around 72,000 putative NCLDV MAGs.

Filtering of GVMAGs

Contigs with a length of less than 5 kb were removed from GVMAGs. Filtering was performed on the basis of the copy number of NCVOGs¹⁶ (Supplementary Tables 2, 3). GVMAGs were removed when they encoded more than 20 copies of NCVOG0023, 4 copies of NCVOG0038, 12 copies of NCVOG0076, 7 copies of NCVOG0249 or 4 copies of NCVOG0262. On the basis of the copy numbers of 16 conserved NCVOGs (NCVOG0035, NCVOG0036, NCVOG0038, NCVOG0052, NCVOG0059, NCVOG0211, NCVOG0249, NCVOG0256, NCVOG0262, NCVOG1060, NCVOG1088, NCVOG1115, NCVOG1117, NCVOG1122, NCVOG1127 and NCVOG1192), which are usually present at low copy numbers across all published NCLDV genomes, a duplication ratio was calculated as follows. The total number of copies of the 16 NCVOGs in the respective GVMAG was divided by the total number of unique observations of the 16 NCVOGs. GVMAGs with a duplication ratio higher than three were excluded from the dataset. We then used Diamond BLASTp⁴⁷ against the NCBI non-redundant (nr) database (August 2018) and assigned a taxonomic affiliation on the basis of best BLASTp hits against Archaea, Bacteria, Eukaryota, phages or other viruses (including NCLDVs) to proteins using an *E*-value cut-off of 1×10^{-5} . Best hits of query proteins to proteins derived from MAGs from the *Tara* Mediterranean metagenome binning survey⁵⁷ were disregarded owing to the high number of misclassified genomes in this dataset. Proteins without a hit in the NCBI nr database were labelled as 'Unknown'. We then applied filters to remove contigs from GVMAGs on the basis of the distribution of taxonomic affiliation of best blast hits (Supplementary Table 7). Finally, alignments were built with mafft⁴⁸ (v.7.294b) for NCVOG0023, NCVOG0038, NCVOG0076, NCVOG0249 and NCVOG0262. Positions with 90% or more gaps were removed from the alignments with trimal⁵⁸ (v.1.4). Protein alignments were concatenated and a species tree constructed with IQ-tree⁵⁹ (LG + F + R8, v.1.6.10). The phylogenetic tree was then manually inspected and for each clade outliers were removed on the basis of the presence, absence and copy numbers of 20 conserved NCVOGs¹⁶, duplication factor (see above), coding density, GC content and genome size. In addition, GVMAGs that represented singletons on long branches were manually removed. The filtered dataset was then clustered together with all available NCLDV reference genomes (December 2018) using average nucleotide identities of greater than 95% and an alignment fraction of at least 50% with FastANI⁶⁰ (v.1.1). For each 95% average nucleotide identity cluster the 6 NCVOGs¹⁶ with the on-average longest amino acid sequences (NCVOG0022, NCVOG0023, NCVOG0038, NCVOG0059, NCVOG0256 and NCVOG1117) were subjected to a within-cluster all-versus-all BLASTp. GVMAGs that had any full-length 100% identity hits between any of these marker proteins to other cluster members were removed from the dataset as potential duplicates. Duplicate GVMAGs originating from the conventional binning approach were removed first and GVMAGs with the largest assembly size were retained.

GVMAG quality on the basis of estimated completeness and contamination

Estimation of the quality of MAGs is critical for their interpretation and use in downstream applications. Standards exist for bacterial and archaeal MAGs that have proposed a three-tier classification (high, medium or low quality) based on estimated genome completeness and contamination⁵⁵. These completeness and contamination metrics are typically calculated on the basis of a set of universal single-copy marker genes. A set of conserved genes in the NCLDV are the NCVOGs¹⁶, of which a subset has been shown to be probably vertically inherited¹⁶ (NCVOG20, Supplementary Table 2). We calculated for each superclade the average number of NCVOG20 present either as a single copy or as multiple copies (Supplementary Table 3). We then compared the number of observed single- and multicopy NCVOG20 in every GVMAG to the mean number of observations in the respective superclade. Considering the high genome plasticity of NCLDVs^{2,61}, we tolerated a deviation from the mean by a factor of 1.2, which was considered low contamination, and a factor of 2 was considered medium contamination (Extended Data Fig. 4 and Supplementary Table 4). Higher deviations from the superclade mean were potentially caused by a non-clonal composition of the GVMAG; these were, as a consequence, considered to be of high contamination. We also estimated completeness on the basis of the presence of the NCVOG20 compared with other members of the respective superclade. The presence of 90% or more of the NCVOG20 compared with the superclade mean resulted in a classification as high quality in terms of completeness. If at least 50% of NCVOG20 were present in a GVMAG then the respective GVMAG was classified as medium quality in terms of estimated completeness, or low if less than 50% of NCVOG20 were present (Extended Data Fig. 4 and Supplementary Table 4). The final GVMAG quality was determined on the basis of a combination of contamination and completeness (Supplementary Table 8). Additional criteria to assign GVMAGs to the high-quality category were the presence of no more than 30 contigs, a minimum assembly size of 100 kb and the presence of at least one contig with a length greater than 30 kb. To assign a GVMAG to the medium-quality category were the presence no more than 50 contigs, a minimum assembly size of 100 kb and the presence of at least one contig with a length greater than 15 kb.

Annotation of GVMAGs

Gene calling was performed with GeneMarkS using the virus model⁶². For functional annotation proteins were subject to BLASTp against previously established NCVOGs¹⁶ and the NCBI nr database (May 2019) using Diamond (v.0.9.21) BLASTp⁴⁷ with an *E*-value cut-off of 1.0×10^{-5} . In addition, protein domains were identified by pfam_scan.pl (v.1.6) against Pfam-A⁶³ (v.29.0), and rRNAs and introns were identified with cmsearch using the Infernal package⁶⁴ (v.1.1.1) against the Rfam database⁶⁵ (v.13.0). No rRNA genes were detected in the final set of GVMAGs. The eggNOG mapper⁶⁶ (v.1.0.3) was used to assign functional categories to NCLDV proteins. Protein families were inferred with PorthoMCL⁶⁷ (version of December 2018) with default settings.

Survey of the NCLDV MCP

We used hmmsearch (v.3.1b2, <http://hmmer.org/>) optimized for the supercomputer Cori⁴⁴ to identify all copies of MCP encoded in the final set of GVMAGs and NCLDV reference genomes. Proteins were extracted and multiple sequence alignments were created with mafft⁴⁸ (v.7.294b) for 74 NCLDV lineages with at least 5 copies of MCP. For each lineage-specific MCP alignment, we inferred models with hmmbuild (v.3.1b2, <http://hmmer.org/>). Using these models, the modified version of hmmsearch (v.3.1b2, <http://hmmer.org/>)⁴⁴ was used to identify all MCPs in the entire set of metagenomes (IMG/M⁴³, June 2018), MCPs with identical amino acid sequences were excluded as potential duplicates. A logistic-regression-based classifier (sklearn LogisticRegression,

solver = 'lbfgs', multi_class = 'ovr') was trained for each NCLDV lineage taking into account the score distribution of all lineage MCPs hits against the entire set of lineage-specific MCP models. The accuracy of the classifier was 0.861. Unbinned metagenomic MCPs were assigned to NCLDV lineages if the classifier returned a probability greater than 50% (sklearn predict_proba), or as 'novel' if the probability was 50% or below. We then normalized the environmental MCP counts on the basis of the observed average copy number of MCP in GVMAGs and reference genomes in the respective lineage. Distribution of NCLDV lineages on the basis of MCPs was projected on a world map with Python 3/basemap on the basis of coordinates provided in IMG metagenomes⁴³.

NCLDV species tree

To build a species tree of the extended NCLDV, viral genomes with at least three out of five core NCVOGs¹⁶ were selected: DNA polymerase elongation subunit family B (NCVOG0038), D5-like helicase-primase (NCVOG0023), packaging ATPase (NCVOG0249), DNA or RNA helicases of superfamily II (NCVOG0076), and poxvirus late transcription factor VLTf3-like (NCVOG0262). The NCVOGs were identified with hmmsearch (version 3.1b2, <http://hmmsearch.org/>) using an *E*-value cut-off of 1×10^{-10} , extracted and aligned using mafft⁴⁸ (v.7.294b). Columns with less than 10% sequence information were removed from the alignment with trimal⁵⁸. The species tree was then calculated on the basis of the concatenated alignment of all five proteins with IQ-tree⁵⁹ (v.1.6.10) with ultrafast bootstrap⁶⁸ and LG + F + R8 as suggested by model test as the best-fit substitution model⁶⁹. The percentage increase in phylogenetic diversity⁷⁰ was calculated on the basis of the difference of the sum of branch lengths of the phylogenetic species trees of the NCLDV including the GVMAGs compared with a NCLDV species tree calculated from published NCLDV reference genomes ($n = 205$, no dereplication based on the average nucleotide identity) with IQ-tree as described above. Phylogenetic trees were visualized with iTol⁷¹ (v.5). Genus or subfamily level lineages were defined on the basis of their monophyly in the species tree and presence or absence pattern of conserved NCVOGs (Supplementary Table 4). If no viral isolates were present in the respective monophyletic clade we designated it MGVL. Neighbouring lineages with isolates and MGVLs were further combined under the working term superclade. Branch lengths separating clades differ based on the density of sampled viruses.

Protein trees

Target proteins were extracted from NCLDV genomes and used to query the NCBI nr database (June 2018) with Diamond BLASTp⁴⁷. The top-50 hits per query were extracted, merged with queries, dereplicated on the basis of protein accession number and aligned with MAFFT (-linsi, v.7.294b)⁴⁸, trimmed with trimal⁵⁸ (removal of positions with more than 90% of gaps) and maximum-likelihood phylogenetic trees inferred with IQ-tree⁵⁹ (multicore v.1.6.10) using ultrafast bootstrap⁶⁸ and the model suggested by the model test feature implemented in IQ-tree⁶⁹ based on Bayesian information criterion. Selected models are indicated in the legend of Extended Data Fig. 8. Owing to its size, the phylogenetic tree for ABC transporter was inferred with FastTree⁷² (v.2.1.10) LG and can be accessed at <https://bitbucket.org/berkeleylab/mtg-gv-exp/>. Phylogenetic trees were visualized with iTol⁷¹ (v.5). Information on functional genes including parent contigs is provided in Supplementary Table 5.

Virus–host linkage through HGT

To generate a cellular nr database, all non-cellular sequences and sequences from the *Tara* Mediterranean genome study⁵⁷ were removed from the NCBI nr database. All proteins in the NCLDV genomes were then subjected to Diamond BLASTp⁴⁷ against the cellular nr database using an *E*-value cut-off of 1×10^{-50} , an alignment fraction of 50% and a minimum sequence identity of 50%. Best blast hits within the same lineage were removed. Proteins that had a hit in cellular nr with a lower

E value compared with hits in the NCLDV blast database were considered HGT candidates. The total number of best hits from lineage pan-proteomes against defined groups of Eukaryotes were then used as edge weights to build an HGT network. The network was created in Gephi (v.0.92)⁷³ using a force layout and filtered at an edge weight of 2. Pfam annotations of HGT candidates were based on the most commonly detected domains and functional categories were assigned with the eggNOG Mapper (v.1.03)⁶⁶. Information on HGT candidates including parent contigs is provided in Supplementary Table 6. The number of HGT linkages was limited by the available of reference genomes and the stringency applied.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

All GVMAGs of estimated high and medium quality with an N50 of greater than 50 kb and estimated low contamination have been deposited at NCBI GenBank as MN738741–MN741037 under BioProject ID PRJNA588800. Nucleotide and protein sequences of GVMAGs can be directly downloaded from <https://genome.jgi.doe.gov/portal/GVMAGs> and <https://figshare.com/s/14788165283d65466732>, and will be available in the Integrated Microbial Genome/Virus (IMG/VR) system⁷⁴ at time of the v.3.0 release. All of the sequence data and metadata from the samples used in this study can further be accessed through the IMG/M system⁴³ (<https://img.jgi.doe.gov>) and NCBI SRA using the metagenome identifiers provided in Supplementary Table 1. Sequence alignments, phylogenetic trees and other data underlying this study can be downloaded from <https://genome.jgi.doe.gov/portal/GVMAGs>.

Code availability

The NCLDV classifier can be obtained from <https://bitbucket.org/berkeleylab/mtg-gv-exp/>.

43. Chen, I. A. et al. IMG/M: integrated genome and metagenome comparative data analysis system. *Nucleic Acids Res.* **45**, D507–D516 (2017).
44. Arndt, W. Modifying HMMER3 to run efficiently on the Cori supercomputer using OpenMP tasking. In *Proc. 2018 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)* 239–246 (2018).
45. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).
46. Emms, D. M. & Kelly, S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **16**, 157 (2015).
47. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).
48. Katoh, K. & Standley, D. M. A simple method to control over-alignment in the MAFFT multiple sequence alignment program. *Bioinformatics* **32**, 1933–1942 (2016).
49. Hyatt, D. et al. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).
50. Liaw, A. & Wiener, M. Classification and regression by randomForest. *R News* **2**, 18–22 (2002).
51. R Core Team. *R: A Language and Environment for Statistical Computing* <http://www.R-project.org/> (R Foundation for Statistical Computing, 2013). (2013).
52. Torchiano, M. effsize: efficient effect size computation. R package version 0.5.4 <https://cran.r-project.org/web/packages/effsize/effsize.pdf> (2015).
53. Kang, D. D., Froula, J., Egan, R. & Wang, Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3**, e1165 (2015).
54. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
55. Bowers, R. M. et al. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* **35**, 725–731 (2017).
56. Kang, D. D. et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **7**, e7359 (2019).
57. Tully, B. J., Sachdeva, R., Graham, E. D. & Heidelberg, J. F. 290 metagenome-assembled genomes from the Mediterranean Sea: a resource for marine microbiology. *PeerJ* **5**, e3558 (2017).

58. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
59. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
60. Jain, C., Rodríguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **9**, 5114 (2018).
61. Filée, J. Route of NCLDV evolution: the genomic accordion. *Curr. Opin. Virol.* **3**, 595–599 (2013).
62. Borodovsky, M. & Lomsadze, A. Gene identification in prokaryotic genomes, phages, metagenomes, and EST sequences with GeneMarkS suite. *Curr. Protoc. Bioinformatics* **35**, 4.5.1–4.5.17 (2011).
63. Finn, R. D. et al. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* **44**, D279–D285 (2016).
64. Nawrocki, E. P. & Eddy, S. R. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **29**, 2933–2935 (2013).
65. Kalvari, I. et al. Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Res.* **46**, D335–D342 (2018).
66. Huerta-Cepas, J. et al. Fast genome-wide functional annotation through orthology assignment by eggNOG-Mapper. *Mol. Biol. Evol.* **34**, 2115–2122 (2017).
67. Tabari, E. & Su, Z. PorthoMCL: parallel orthology prediction using MCL for the realm of massive genome availability. *Big Data Analytics* **2**, 4 (2017).
68. Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
69. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
70. Wu, D. et al. A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature* **462**, 1056–1060 (2009).
71. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–W245 (2016).
72. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).
73. Bastian, M., Heymann, S. & Jacomy, M. Gephi: an open source software for exploring and manipulating networks. *Proc. International AAAI Conference on Weblogs and Social Media* (2009).
74. Paez-Espino, D. et al. IMG/VR v2.0: an integrated data management and analysis system for cultivated and environmental viral genomes. *Nucleic Acids Res.* **47**, D678–D686 (2019).

Acknowledgements This work was conducted by the US Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, under contract no. DE-AC02-05CH11231 and made use of resources of the National Energy Research Scientific Computing Center, which is also supported by the DOE Office of Science under contract no. DE-AC02-05CH11231. We thank the DOE JGI user community and Tara Oceans for sampling efforts and for providing the metagenomic data that underlies this study and E. Kirton for running CheckM for non-targeted binning of the IMG/M metagenomes.

Author contributions F.S. and T.W. conceived the study. D.A.W., V.J.D., K.D.M. and K.T.K. provided metagenomic datasets with a large number of GVMAGs. F.S. performed targeted binning of public metagenomes, phylogenomics, analysis of functional genes and HGT analysis. S.R. developed and benchmarked RBS/gene density classifier. D.P.-E. provided initial HMMs for the NCLDV major capsid protein. S.J. performed non-targeted binning of public metagenomes in IMG. F.S. and S.R. performed quality control of GVMAGs. F.S. visualized the data. T.W., N.C.K. and E.A.E.-F. supervised research. F.S. and T.W. prepared the manuscript, with contributions from all authors. All authors read and approved the final manuscript.

Competing interests The authors declare no competing interests.

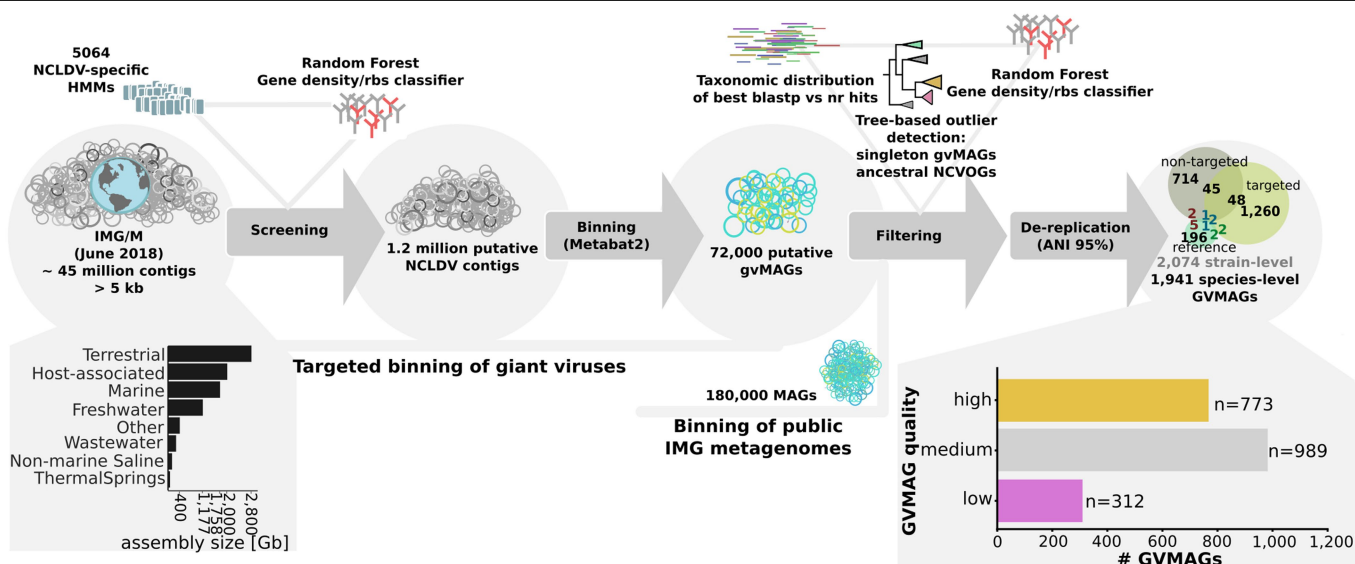
Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-1957-x>.

Correspondence and requests for materials should be addressed to F.S. or T.W.

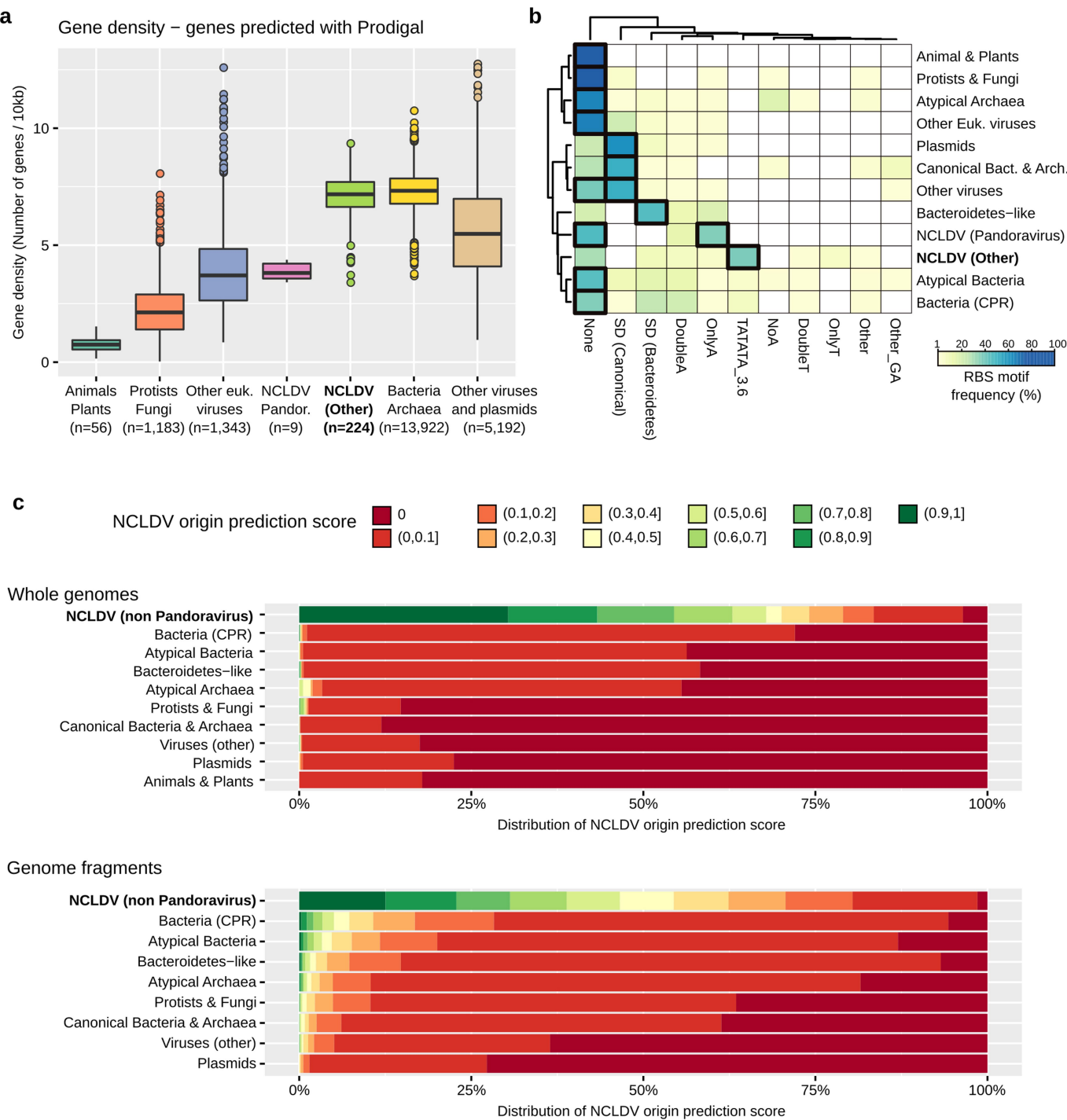
Peer review information *Nature* thanks Hisashi Endo, Mart Krupovic and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



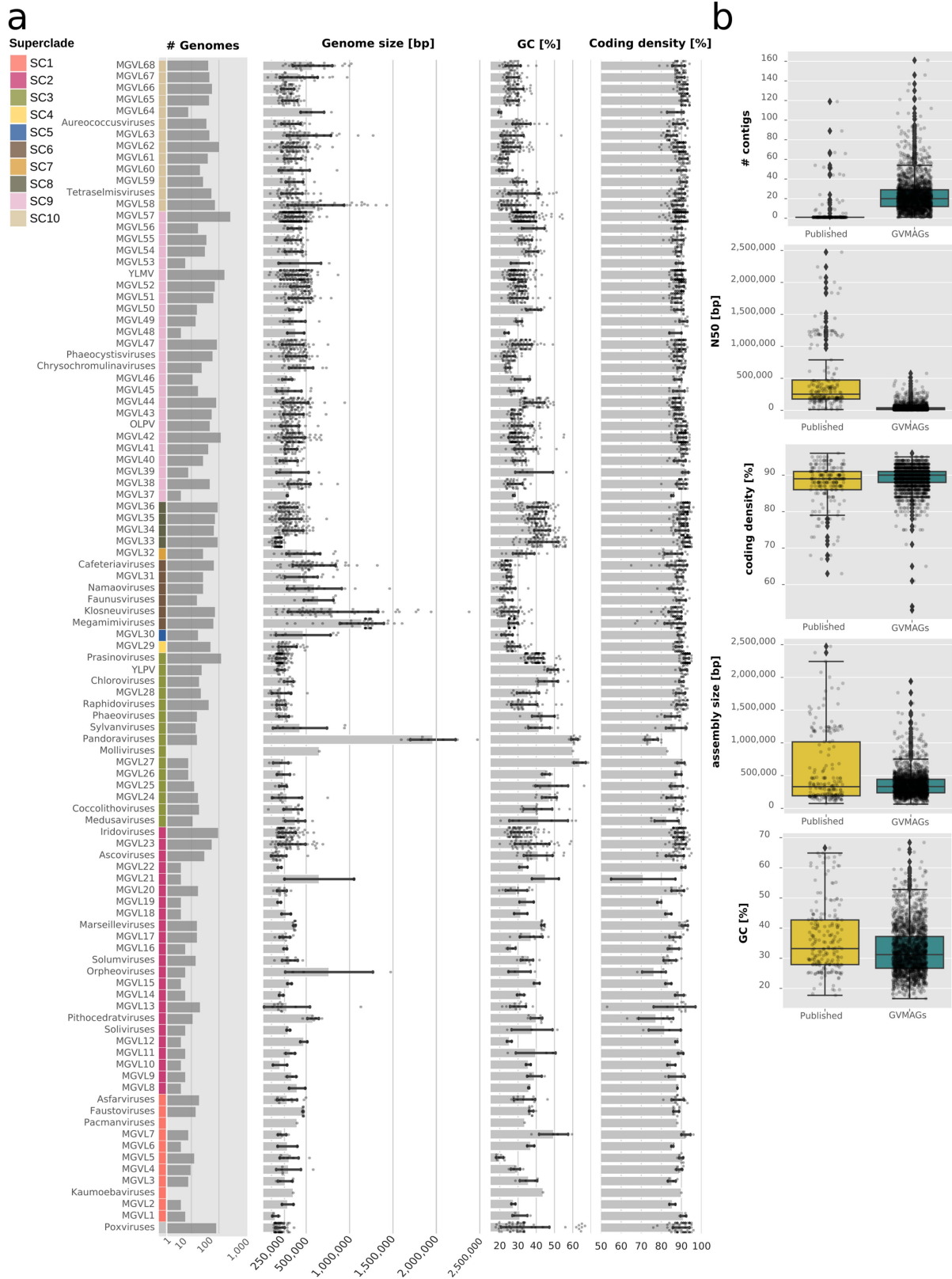
Extended Data Fig. 1 | Discovery pipeline for GVMAGs. Approximately 46 million contigs that were longer than 5 kb and were available in IMG/M¹⁵ (June 2018) were screened for potential NCLDV contigs using a combination of 5,064 NCLDV-specific HMMs and a random-forest classifier based on gene density and RBS motifs. The resulting set of 1.2 million contigs was then subjected to metagenomic binning using MetaBAT2⁵³, with binning performed separately for each metagenome that contained putative NCLDV contigs. To the resulting approximately 72,000 GVMAGs, we added around 180,000 low-quality MAGs based on MIMAG⁵⁵ that were generated by non-targeted binning

of metagenomes in IMG/M. The resulting set of approximately 252,000 GVMAGs and MAGs were then filtered on the basis of assembly size and using a combination of the consensus of taxonomic affiliation of best blast hits across contigs, the presence or absence and copy numbers of frequently conserved NCLDV genes taking into account neighbouring taxa in the species tree and random-forest classifier based on gene density and RBS motifs. Outlier contigs were removed as described in the Methods and only MAGs that showed a copy-number distribution of frequently conserved NCLDV genes similar to closely related viral genomes were maintained in the final dataset.



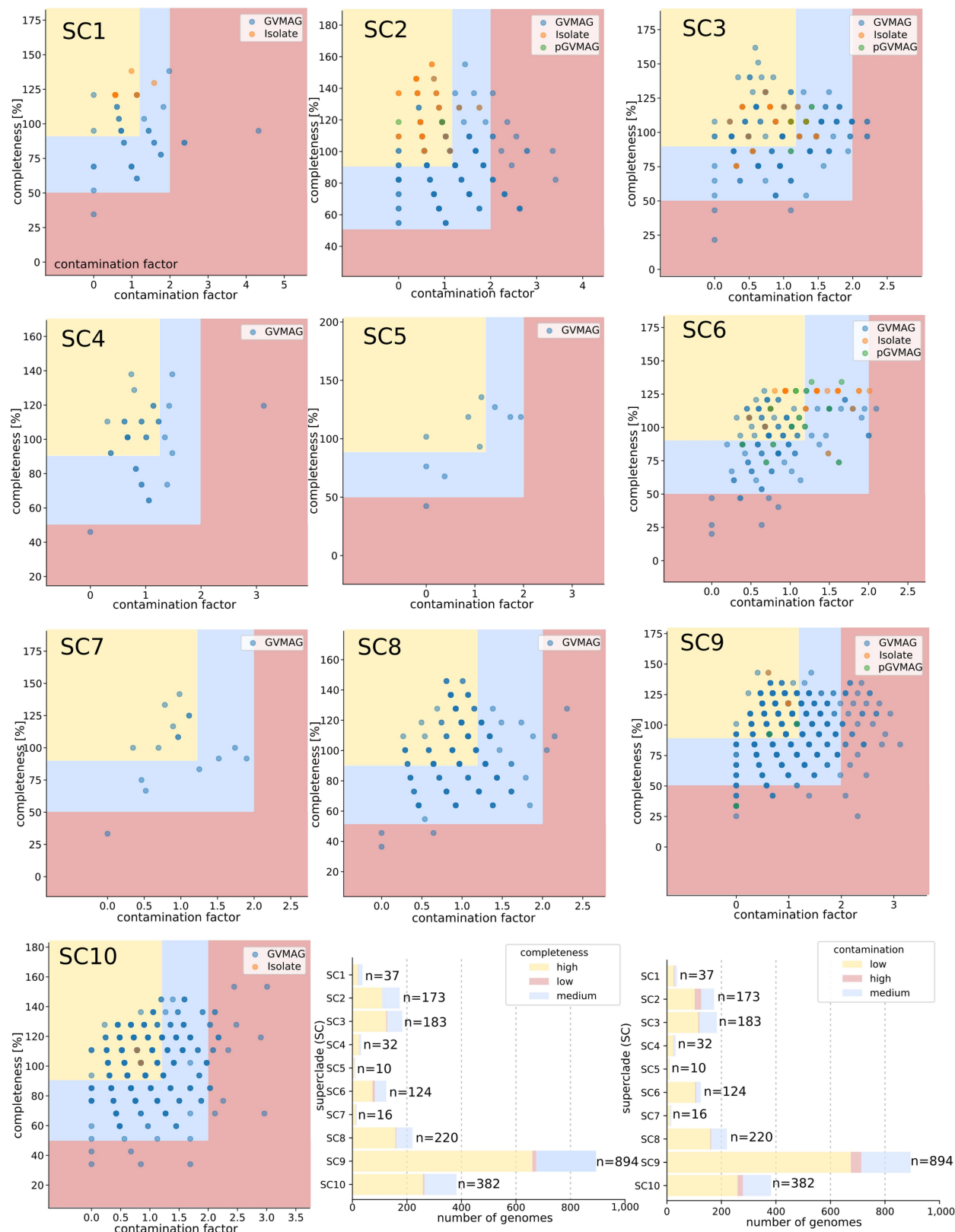
Extended Data Fig. 2 | The RBS classifier. Unique features of NCLDV genomes and efficiency of random-forest classifiers based on these features. **a**, Gene density (y axis, average number of genes predicted per 10 kb of genome) for genomic sequences from different types of organisms or entities (x axis). Genomes were grouped on the basis of taxonomy (kingdom and domain ranks) as well as patterns of RBS motifs and gene density. ‘Other euk. viruses’, non-NCLDV eukaryotic viruses; ‘NCLDV Pandor.’, pandoravirus and similar NCLDVs; ‘NCLDV (Other)’, non-pandoravirus NCLDVs. Centre lines of box plots represent the median, bounds of the boxes indicate the lower and upper quartiles, whiskers extend to points that lie within 1.5× the interquartile range of the lower and upper quartiles. Sample sizes (number of genomes) are indicated. **b**, Frequency of RBS motifs identified across different genomes

groups. RBS motif frequencies were based on prodigal gene prediction using the ‘full motif scan’ option. For clarity, only RBS motif frequencies >1% are displayed. RBS motif frequencies ≥30% are highlighted with a bold outline. ‘Other Euk. viruses’, non-NCLDV eukaryotic viruses; ‘NCLDV (pandoravirus)’, pandoravirus and similar NCLDVs; ‘NCLDV (Other)’, non-pandoravirus NCLDVs. **c**, Predictions of NCLDV origin on the basis of genome features and predicted RBS motifs by random-forest classifiers for complete genomes (top) and short genome fragments (bottom). Predictions for individual genomes were obtained through a tenfold cross-validation. Similar results were obtained when predicting only two classes (NCLDV and non-NCLDV, displayed here) or when predicting classes corresponding to the eight types of genomes. CPR, candidate phyla radiation; SD, Shine–Dalgarno sequence.



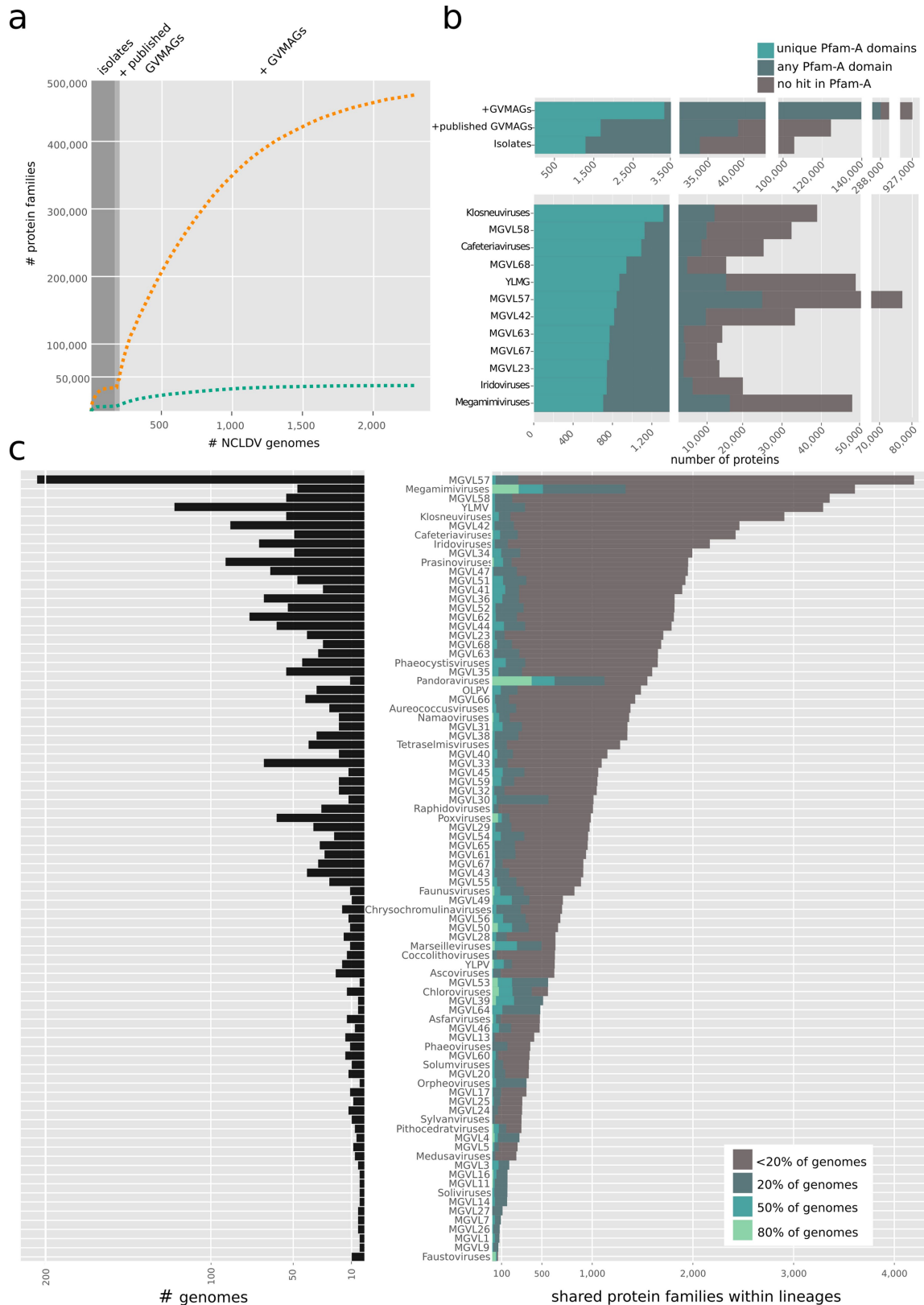
Extended Data Fig. 3 | Features of GVMAGs. **a**, Mean assembly size, GC content and coding density for each lineage in the NCLDV, coloured by superclade, individual data points are shown. Data are mean \pm s.d. **b**, Assembly metrics of all GVMAGs compared to previously published NCLDV genomes included in this

study. Centre lines of box plots represent the median, bounds of boxes indicate the lower and upper quartiles, whiskers extend to points that lie within 1.5 \times interquartile range of the lower and upper quartiles. Sample size for the published data is 205 genomes and for GVMAGs is 2,074 genomes.



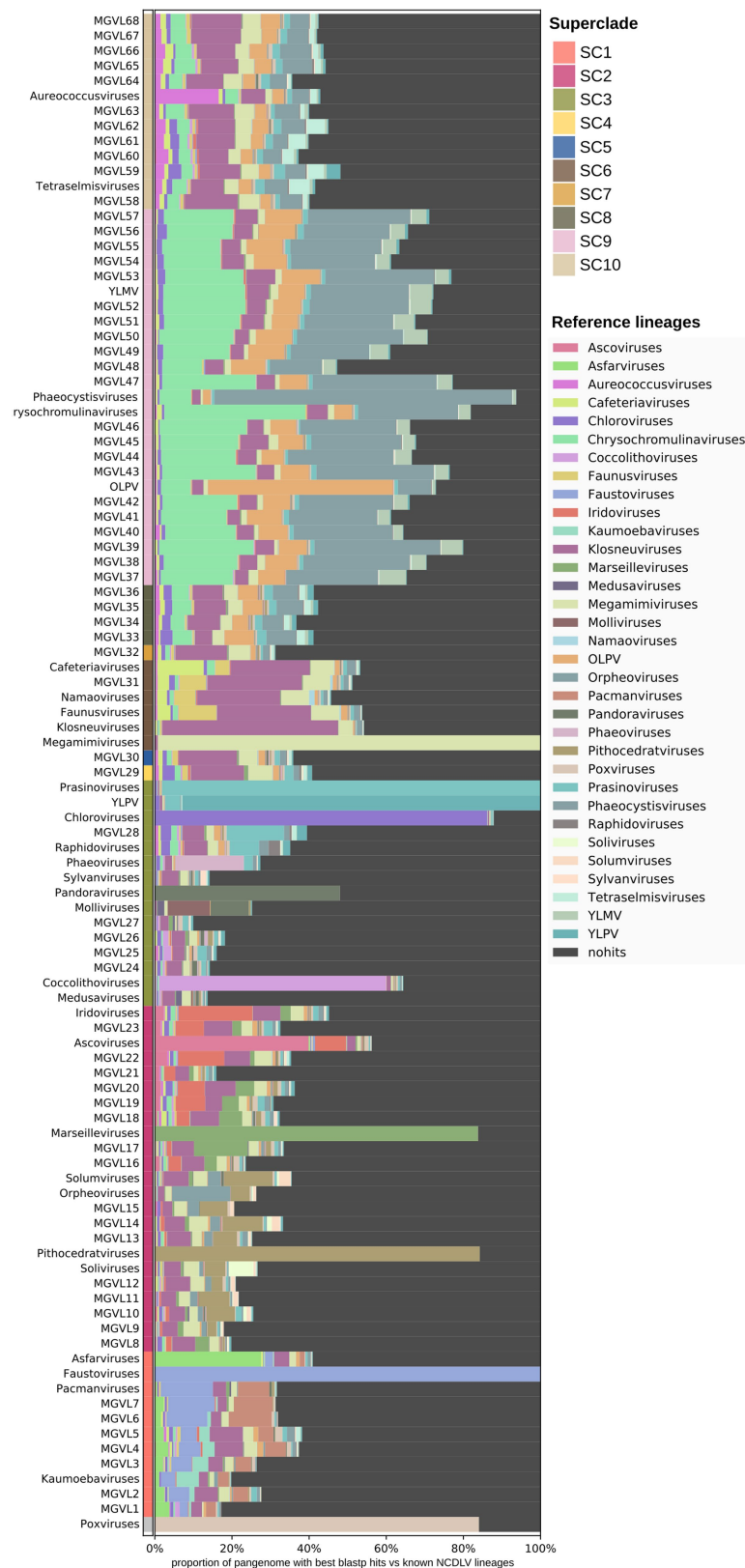
Extended Data Fig. 4 | Estimated completeness and contamination of GVMAGs on the basis of the presence of conserved NCVOGs. Scatter plots show estimated completeness and contamination for GVMAGs in each superclade (SC), previously published GVMAGs (pGVMAGs) and isolate genomes (filled circles with different colours) compared with the average of the respective superclade. Genomes in the red area were classified as low

quality, genomes in the blue area were classified as medium quality and genomes in the yellow area were classified as high quality on the basis of the combination of completeness and contamination. Stacked bars (bottom right) summarize, for each NCLDV superclade, the total number of GVMAGs with low, medium and high contamination and completeness.

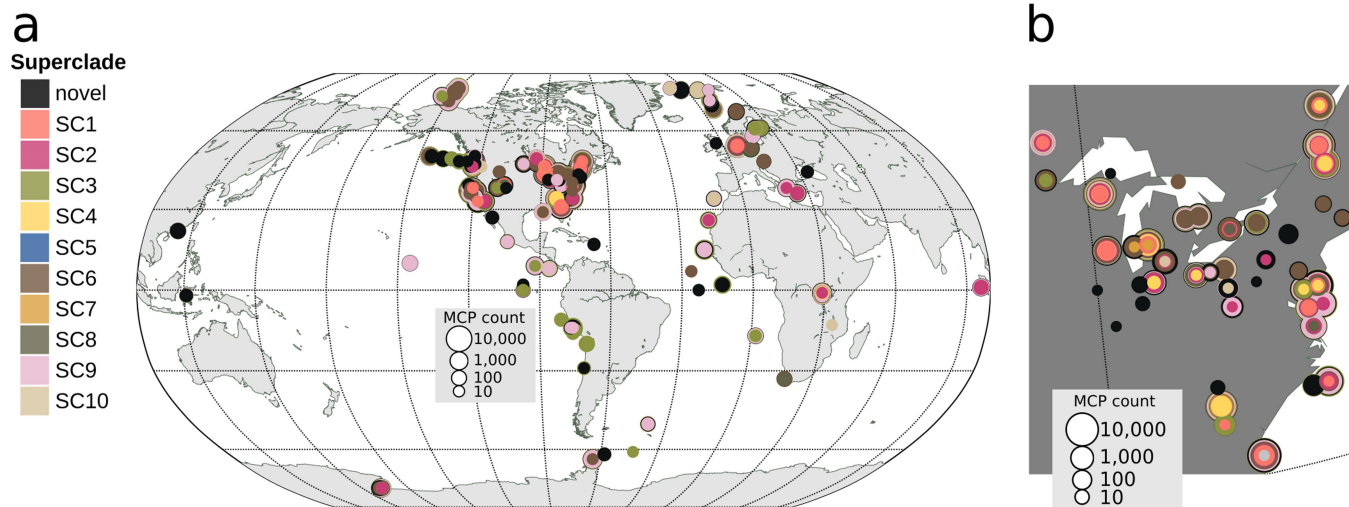


Extended Data Fig. 5 | Shared and unique protein families within NCLDV lineages. **a**, Collector's curve showing the increase in functional diversity estimated on the basis of the total number of protein families detected in NCLDV isolates, previously published GVMAGs and GVMAGs recovered in this study. The orange curve includes all detected protein families; the blue curve only includes protein families that included by at least two proteins. **b**, Top, the total number of different Pfam-A domains, total number of proteins with any

Pfam-A domain and total number of proteins found in NCLDV isolates, previously published NCLDV genomes from metagenomes and GVMAGs recovered in this study. Bottom, NCLDV lineages with the greatest number of unique Pfam-A domains. **c**, The total number of genomes per lineage (left) and total number of protein families (at least two members) found in each lineage are indicated together with the proportion of genomes in the respective lineage that share protein families (right).

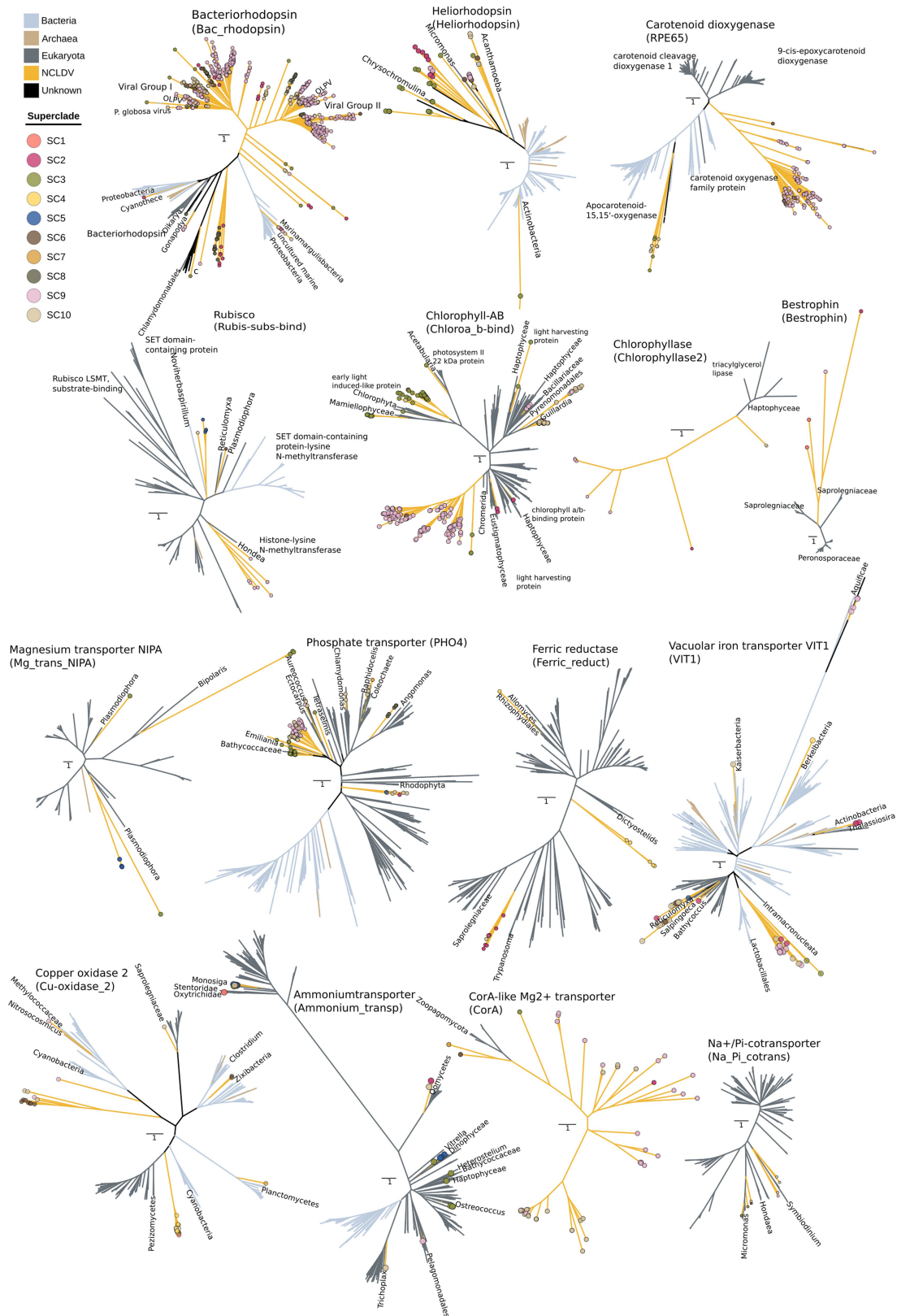


Extended Data Fig. 6 | Similarity of proteins encoded in expanded NCDLV lineages and new MGVLs to known NCDLV proteins. For each lineage the proportion of encoded proteins with homology (E -value cut-off of 1×10^{-5}) to known NCDLV proteins is shown.



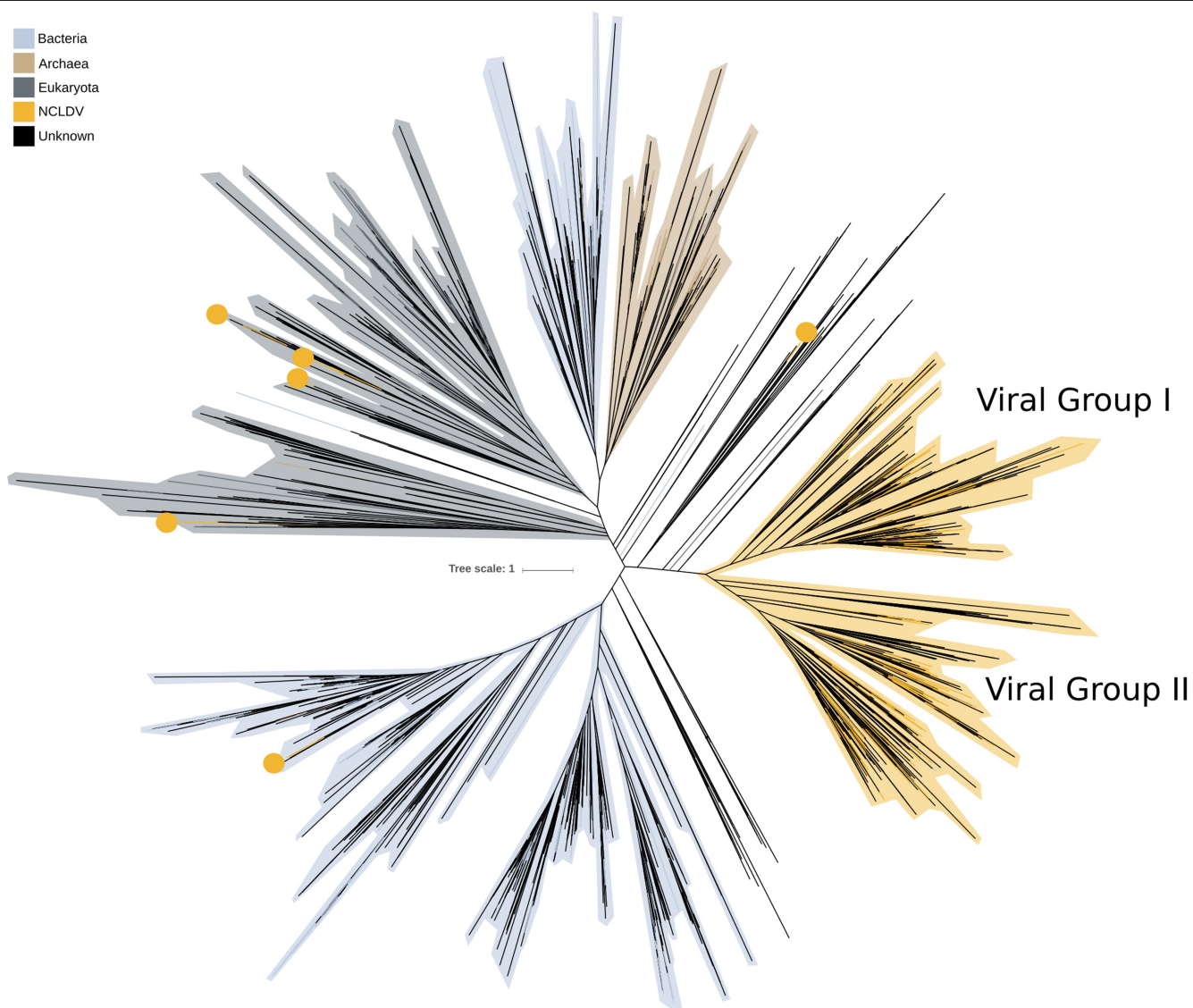
Extended Data Fig. 7 | Distribution of NCLDV MCPs. a, Global distribution of NCLDV MCPs. **b,** A detailed view of the Midwest and East Coast of the United States and Canada. Filled circles are coloured on the basis of the affiliation with superclade and the circle diameter correlates with the number of MCPs

detected at the respective sampling location. Circles at the same coordinates are stacked by size with the largest circles at the bottom. The category 'novel' contains all MCPs that could not be assigned to any of the superclades.



Extended Data Fig. 8 | Maximum-likelihood phylogenetic trees. Maximum-likelihood phylogenetic trees that underlie the analysis in Fig. 2. Trees were inferred using IQ-tree with the following models: Na⁺/P_i cotransporter, LG4M + R7; ammonium transporter, LG4M + R10; bacteriorhodopsin, LG + F + R10; bestrophin, LG4M + R5; carotenoid dioxygenase, LG + F + R10;

Chlorophyll *ab*, LG4M + F + R10; chlorophyllase, LG + I + G4; CorA-like Mg²⁺ transporter, LG + F + R3; copper oxidase II, LG4M + R10; heliorhodopsin, LG4M + R9; magnesium transporter NIPA, LG4M + R6; ferric reductase, LG + F + R9; phosphate transporter, LG4M + R10; Rubisco, LG4M + R6; and vacuolar iron transporter (VIT1), LG4M + R10.



Extended Data Fig. 9 | Diversity of metagenomic rhodopsins. Maximum-likelihood tree (IQ-tree, LG4M + R10 substitution model) of rhodopsins after dereplication through clustering with CD-hit at a 70% similarity threshold. Clades that predominantly include rhodopsins of archaeal, bacterial,

eukaryotic or NCLDV origin are highlighted in the different colours. Yellow filled circles indicate NCLDV rhodopsins that have probably been acquired from cellular organisms through HGT.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | | |
|-------------------------------------|---|
| n/a | Confirmed |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

IMG/M June2018
 NCBI Genbank (June 2018)
 NCBI non redundant database May 2019
 Pfam-A (v 29.0)
 Rfam (v 13.0)

Data analysis

hmmmer (version 3.1b2)
 orthofinder (v2.27)
 mafft (v7.294b)
 Prodigal (v2.6.3)
 MetaBAT (v0.32.4)
 MetaBAT (v2)
 CheckM (v1.0.7)
 trimAL (v1.4)
 IQ-tree (1.6.10)
 FastANI (v1.1)
 Diamond (v0.9.21)
 Infernal (v1.1.1)
 pfam_scan.pl (v1.6)
 iTOL (v5)
 eggNOG mapper (v1.0.3)
 FastTree (v.2.1.10)
 Gephi (v0.92)

Scikit-learn (v0.20.3)
 stats package (v4) in R
 PorthoMCL (version of December 2018)
 NCLDV classifier is available at <https://bitbucket.org/berkeleylab/mtg-gv-exp/>

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All GVMAGs of estimated high and medium quality with an N50 of greater than 50kb and estimated 'low' contamination have been deposited at NCBI Genbank under BioProject ID PRJNA588800. Nucleotide and protein sequences of GVMAGs can be directly downloaded from <https://genome.jgi.doe.gov/portal/GVMAGs> and will become available in IMG/VR74 at time of the v.3.0 release. All the sequence data and metadata from the samples used in this work can further be accessed through the Integrated Microbial Genomes and Microbiomes (IMG/M) systems43 (<https://img.jgi.doe.gov>) and NCBI SRA using the metagenome identifiers provided in Supplementary Table 1. Sequence alignments, phylogenetic trees and other data underlying this study can be downloaded from <https://genome.jgi.doe.gov/portal/GVMAGs>

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Recovery of nucleocytoplasmic large DNA virus metagenome assembled genomes from all publicly available metagenome data in IMG (https://img.jgi.doe.gov/)
Research sample	No samples were taken for this study, subject of this study was all publicly available metagenome data in IMG in June 2018 encompassing 8,535 datasets (https://img.jgi.doe.gov/)
Sampling strategy	No samples were taken for this study, all existing publicly available metagenome data in the IMG/M database (https://img.jgi.doe.gov/) in June 2018 was used in this study.
Data collection	The data was collected in June 2018 from the IMG/M database (https://img.jgi.doe.gov/) by Frederik Schulz
Timing and spatial scale	Metagenomic data was generated between 2008 and 2018 by the DOE JGI User Community and Tara Oceans
Data exclusions	For unpublished metagenome datasets used in this study, PIs are either included as co-authors, or PIs were asked for permission and if permission was denied the datasets were excluded from the analysis
Reproducibility	This experiment has not been reproduced but can be reproduced with the methods outlined in the manuscript.
Randomization	Randomization is not relevant in this study as all available data has been mined for nucleocytoplasmic large DNA virus metagenome assembled genomes
Blinding	Blinding was not relevant for this study as the same methods have been applied to the entire data set.
Did the study involve field work?	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

AQP5 enriches for stem cells and cancer origins in the distal stomach

<https://doi.org/10.1038/s41586-020-1973-x>

Received: 10 November 2018

Accepted: 9 December 2019

Published online: 5 February 2020

Si Hui Tan^{1,18}, Yada Swathi^{1,18}, Shawna Tan¹, Jasmine Goh¹, Ryo Seishima¹, Kazuhiro Murakami², Masanobu Oshima³, Toshikatsu Tsuji⁴, Phyllis Phuah¹, Liang Thing Tan¹, Esther Wong¹, Aliya Fatehullah¹, Taotao Sheng^{5,6}, Shamaine Wei Ting Ho^{5,7}, Heike I. Grabsch^{8,9}, Supriya Srivastava^{10,11}, Ming Teh¹⁰, Simon L. I. J. Denil¹², Seri Mustafah¹³, Patrick Tan^{5,14}, Asim Shabbir¹⁵, Jimmy So¹⁵, Khay Guan Yeoh¹⁶ & Nick Barker^{1,2,17*}

LGR5 marks resident adult epithelial stem cells at the gland base in the mouse pyloric stomach¹, but the identity of the equivalent human stem cell population remains unknown owing to a lack of surface markers that facilitate its prospective isolation and validation. In mouse models of intestinal cancer, LGR5⁺ intestinal stem cells are major sources of cancer following hyperactivation of the WNT pathway². However, the contribution of pyloric LGR5⁺ stem cells to gastric cancer following dysregulation of the WNT pathway—a frequent event in gastric cancer in humans³—is unknown. Here we use comparative profiling of LGR5⁺ stem cell populations along the mouse gastrointestinal tract to identify, and then functionally validate, the membrane protein AQP5 as a marker that enriches for mouse and human adult pyloric stem cells. We show that stem cells within the AQP5⁺ compartment are a source of WNT-driven, invasive gastric cancer in vivo, using newly generated *Aqp5-creERT2* mouse models. Additionally, tumour-resident AQP5⁺ cells can selectively initiate organoid growth in vitro, which indicates that this population contains potential cancer stem cells. In humans, AQP5 is frequently expressed in primary intestinal and diffuse subtypes of gastric cancer (and in metastases of these subtypes), and often displays altered cellular localization compared with healthy tissue. These newly identified markers and mouse models will be an invaluable resource for deciphering the early formation of gastric cancer, and for isolating and characterizing human-stomach stem cells as a prerequisite for harnessing the regenerative-medicine potential of these cells in the clinic.

To identify markers specific to LGR5^{high} pyloric stem cells within the gastrointestinal tract, we profiled the transcriptomes of quantitative (q)PCR-validated LGR5-enhanced green fluorescent protein (eGFP)^{high} (LGR5⁺ stem cells), LGR5-eGFP^{low} (immediate progeny) and unfractionated populations from the small intestine, colon and gastric pylorus of *Lgr5-eGFP-IRES-creERT2* mice⁴ by microarray, and identified genes that are selectively enriched in LGR5-eGFP^{high} pyloric stem cells (Fig. 1a, Extended Data Fig. 1a–d, Supplementary Table 1). This dataset also revealed the transcriptional signature of LGR5-eGFP^{high} colon stem cells (Supplementary Table 2). The profiling of LGR5-eGFP^{high}, LGR5-eGFP^{low} and LGR5-eGFP⁺ pylorus populations from another LGR5 reporter model (*LGR5-DTR-eGFP*, ref. ⁵) revealed 67 overlapping genes (Fig. 1b).

Candidate markers were empirically validated by qPCR, in situ hybridization (ISH) and immunostaining. Optimal candidates presented enriched expression in the LGR5-eGFP^{high} pyloric population by qPCR, robust localized mRNA and/or protein expression within the LGR5⁺ pyloric gland base¹, minimal expression in intestines and gastric corpus and co-expression with *Lgr5* (Extended Data Fig. 1j–n). Six candidates were selected: α-1,4-N-acetylglucosaminyltransferase (A4GNT), aquaporin 5 (AQP5), gastric intrinsic factor (GIF), mucin 6 (MUC6), solute carrier protein 9a3 (SLC9A3 (also known as NHE3)) and secreted phosphoprotein 1 (SPPI, also known as osteopontin) (Fig. 1c–m, Extended Data Fig. 1e–n, Supplementary Table 3). To validate the six candidates as markers of populations enriched in pyloric stem cells, we generated *eGFP-IRES-creERT2* mouse models

¹A*STAR Institute of Medical Biology, Singapore, Singapore. ²Division of Epithelial Stem Cell Biology, Cancer Research Institute, Kanazawa University, Kanazawa, Japan. ³Division of Genetics, Cancer Research Institute, Nano-Life Science Institute, Kanazawa University, Kanazawa, Japan. ⁴Department of Gastroenterological Surgery, Ishikawa Prefectural Central Hospital, Kanazawa, Japan. ⁵Cancer and Stem Cell Biology Program, Duke-NUS Graduate Medical School, Singapore, Singapore. ⁶Department of Biochemistry, National University of Singapore, Singapore, Singapore. ⁷Cancer Science Institute of Singapore, National University of Singapore, Singapore, Singapore. ⁸Department of Pathology, GROW School for Oncology and Developmental Biology, Maastricht University Medical Center+, Maastricht, The Netherlands. ⁹Pathology and Data Analytics, Leeds Institute of Medical Research at St. James's, University of Leeds, Leeds, UK. ¹⁰Department of Pathology, National University Health System, Singapore, Singapore. ¹¹Department of Medicine, National University Health System, Singapore, Singapore. ¹²Skin Research Institute of Singapore, Singapore, Singapore. ¹³A*STAR Singapore Immunology Network, Singapore, Singapore. ¹⁴A*STAR Genome Institute of Singapore, Singapore, Singapore. ¹⁵Department of Surgery, National University Health System, Singapore, Singapore. ¹⁶Department of Medicine, National University of Singapore, Singapore, Singapore. ¹⁷School of Biological Sciences, Nanyang Technological University, Singapore, Singapore. ¹⁸These authors contributed equally: Si Hui Tan, Yada Swathi. *e-mail: Nicholas.barker@imb.a-star.edu.sg

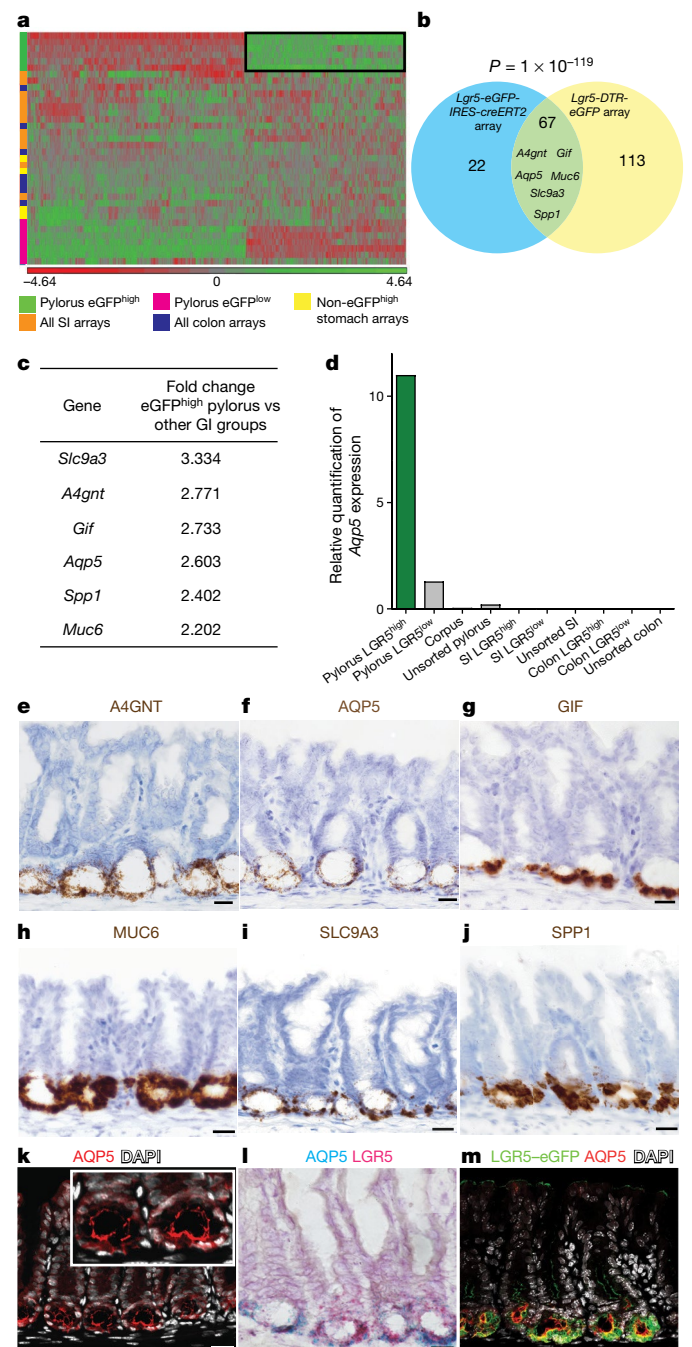


Fig. 1 | Comparative profiling of LGR5 populations in gastrointestinal tissues identifies new pyloric-specific markers. **a**, Heat map of transcriptomes of LGR5-eGFP^{high}, LGR5-eGFP^{low} and unsorted populations from mouse pylorus, small intestine (SI) and colon ($n = 4$, 2 and 3 biological replicates, respectively) from *Lgr5-eGFP-IRES-creERT2* mice. Candidates are enriched only in the GFP^{high} pylorus population (black box). Values are log₂-transformed. **b**, Overlap of candidate markers from *Lgr5-eGFP-IRES-creERT2* (blue) and *Lgr5-DTR-eGFP* (yellow, $n = 4$ biological replicates) models (statistical significance assessed by hypergeometric distribution). **c**, Shortlisted pylorus-specific markers. Fold change is log₂-transformed. GI, gastrointestinal. **d**, Relative *Aqp5* expression (by qPCR) in gastrointestinal populations ($n = 2$ technical replicates of pooled sample from 8 mice, means represented). **e–j**, ISH expression of candidate markers in pylorus ($n = 3$ mice). **k–m**, AQP5 expression with LGR5 in pylorus by immunostaining (**k**, **m**) and ISH (**l**) ($n = 3$ mice). Scale bars, 25 μ m.

for *Aqp5*, *A4gnt*, *Spp1* and *Slc9A3* (Extended Data Fig. 2a), as well as *Aqp5*-2A-eGFP, *Slc9A3*-2A-eGFP, *Aqp5*-2A-creERT2 and *Slc9A3*-2A-creERT2 mice, in which endogenous gene expression is unaffected (Extended Data Fig. 2b). eGFP expression in the pylori of the 2A-eGFP models recapitulated endogenous gene expression in the pylorus and small intestine (Extended Data Fig. 2c–f). Additionally, the 97.4% concurrence between the AQP5⁺ and SLC9A3⁺ pyloric populations revealed by costaining (Extended Data Fig. 2g, h) reaffirms that these markers effectively label the same population.

We evaluated the in vivo contribution of the gland populations that express the candidate genes to epithelial renewal via lineage tracing in *creERT2*; *Rosa26-tdTomato*^{LSL} lines. tdTomato expression was first observed at the gland bases in all lines 20–48 h after tamoxifen induction, confirming the expected Cre expression domain (Extended Data Fig. 3a–d, q–r). After several months (tissue turnover spans 7–10 days¹), multiple glands that were entirely tdTomato⁺ were evident throughout the pylorus, documenting the long-term self-renewal and multipotency of the cells expressing *Aqp5*, *Slc9A3*, *Spp1* or *A4gnt* (Extended Data Fig. 3e–h, s–t). Notably, no intestinal tracing was observed, except for transient reporter expression within the villi of *Slc9A3*-cre models (Extended Data Figs. 2f, 3i–x). These observations confirm that populations that express *Aqp5*, *Spp1*, *Slc9a3* and *A4gnt* contain pyloric stem cells.

AQP5 enriches for stem cells in mice

The AQP5 water channel protein⁶ emerged as a promising candidate for isolating pyloric stem cells from mice and humans. eGFP expression was restricted to pyloric gland bases in *Aqp5-eGFP-IRES-creERT2* and *Aqp5*-2A-eGFP mouse models (Extended Data Fig. 2c, w). Sorted eGFP⁺ cells from adult *Aqp5-eGFP-IRES-creERT2* mice (Fig. 2a) showed a 9-fold and 15-fold enrichment of *Aqp5* and *Lgr5* transcripts, respectively, over eGFP⁺ cells (Extended Data Fig. 2u, v). By immunostaining, endogenous AQP5 protein colocalized with eGFP (Extended Data Fig. 2w). Thus, the AQP5 models faithfully report *Aqp5* expression in pyloric gland bases.

Next, we found that the AQP5⁺ population overlapped with GIF and KI67, but not gastrin (GAST), chromogranin A (CHGA) or mucin 5, subtypes A and C (MUC5AC), whereas the LGR5-eGFP^{high} population expressed major gastric lineage markers (GIF, GAST and CHGA) and KI67 (Fig. 2b–e, Extended Data Fig. 2k–m, x). This observation was confirmed via single-cell analysis of LGR5-eGFP^{high} pyloric stem cells, using CEL-seq and RaceID⁷. The LGR5-eGFP^{high} compartment comprised one major and two minor subpopulations: the major subpopulation co-expressed some of the newly identified pyloric markers (AQP5, GIF and MUC6) (Extended Data Fig. 2n–q), one minor subpopulation expressed CHGA and GAST, and the other expressed KRT8 and KRT18 (Extended Data Fig. 2r, s). Proliferation-marker expression was significantly enriched in the major subpopulation (Extended Data Fig. 2t). AQP5 staining on *Lgr5-DTR-eGFP* pylori revealed a 94.1% overlap between the two populations (Extended Data Fig. 2i, j), underscoring the CEL-seq finding that AQP5 marks the major subpopulation of LGR5^{high} cells.

Lineage tracing using adult *Aqp5-eGFP-IRES-creERT2*; *Rosa26-tdTomato*^{LSL} (*Aqp5-IRES-creERT2*; *tdTomato*) mice detailed the homeostatic behaviour of AQP5⁺ cells: tdTomato⁺ cells appeared exclusively at the gland bases 20 h after tamoxifen induction (Fig. 2f, Extended Data Fig. 3y), expanded to clones reaching gland surfaces by 5 days (Fig. 2g) and persisted for 1 year, demonstrating self-renewal of AQP5⁺ cells (Fig. 2h, i, Extended Data Fig. 3z, c'). Uninduced controls presented negligible tdTomato⁺ clones (Extended Data Fig. 3a', b'). Six months after induction, tdTomato⁺ clones comprised the major pyloric lineages that expressed GIF, GAST, CHGA and MUC5AC (Fig. 2j–m), confirming multipotency in the AQP5⁺ population. Ablating endogenous AQP5⁺ cells using a *Aqp5*-2A-DTR model severely disrupted the gland bases (Extended Data Fig. 4k–r).

Tracing was absent from corpus, small intestine and colon, except for Brunner's glands (Extended Data Fig. 3d'–h'). Our *Aqp5*-2A-eGFP

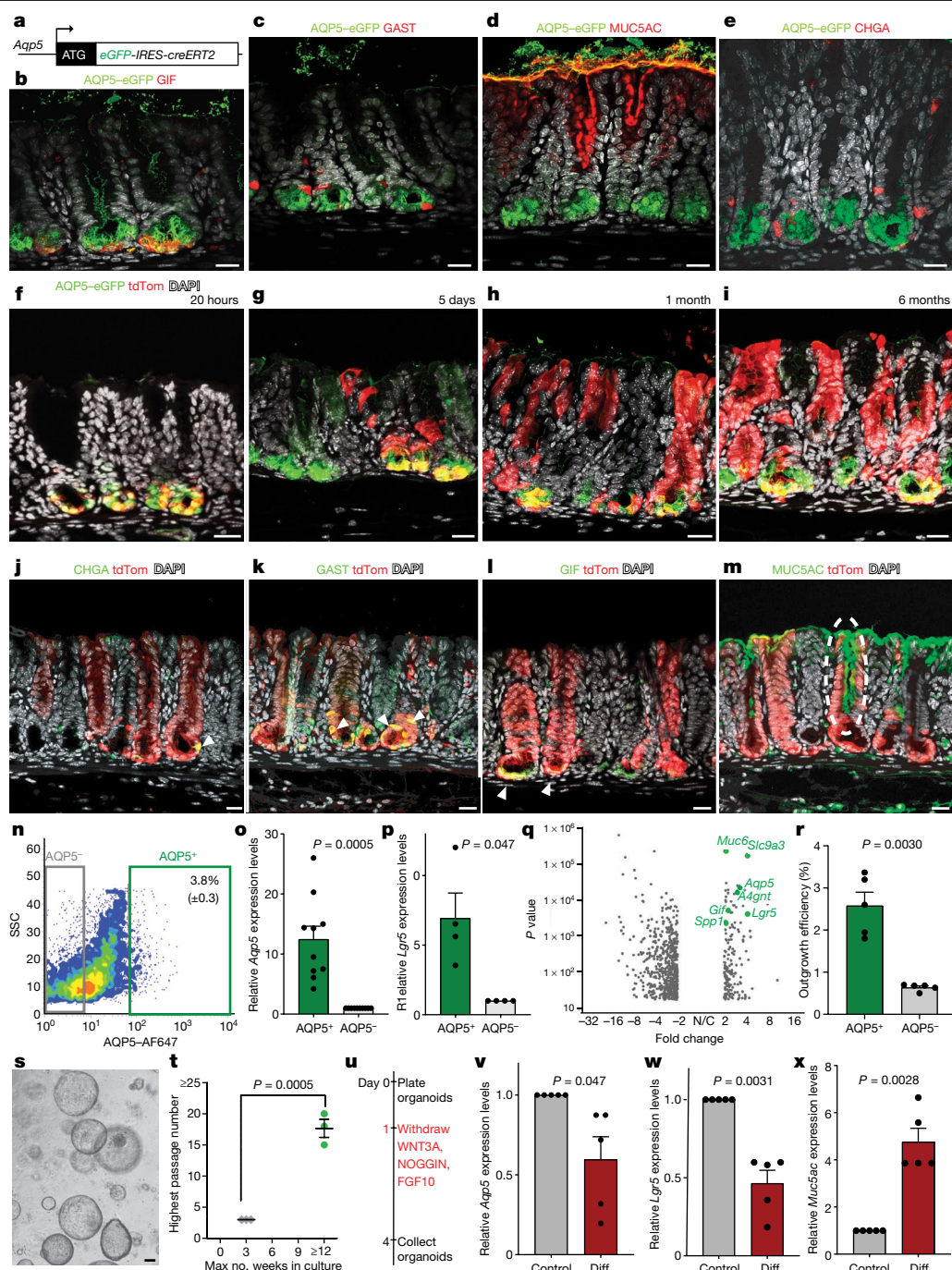


Fig. 2 | Pyloric AQP5⁺ population contains stem cells in vivo. **a**, *Aqp5-eGFP-IRES-creERT2* transgene. **b–e**, eGFP co-immunostaining in *Aqp5-eGFP-IRES-creERT2* pylorus with GIF (**b**), GAST (**c**), MUC5AC (**d**) and CHGA (**e**) ($n = 3$ mice). **f–i**, Lineage tracing of *Aqp5-eGFP-IRES-creERT2;tdTomato^{LSL}* pylori at 20 h (**f**), 5 days (**g**), 1 month (**h**) and 6 months (**i**) ($n = 3$ mice). **j–m**, Co-immunostaining of tdTomato (tdTom) with CHGA (**j**), GAST (**k**), GIF (**l**) and MUC5AC (**m**) ($n = 3$ mice per marker). **n**, FACS gating for sorting AQP5⁺ and AQP5[−] cells. Mean \pm s.e.m., $n = 6$ biological replicates. SSC, side scatter. **o**, **p**, Relative *Aqp5* (**o**) and *Lgr5* (**p**) expression (by qPCR) in sorted populations. Mean \pm s.e.m.; $n = 10$ biological replicates. **q**, *Lgr5* and candidate-marker enrichment in volcano plot of differentially expressed genes in AQP5⁺ population by microarray.

$n = 4$ biological replicates, P value from one-way analysis of variance (ANOVA) in Partek analysis software. N/C, no change. **r–s**, Outgrowth efficiency of single antibody-mediated FACS-sorted AQP5⁺ and AQP5[−] cells (**r**) with representative image (**s**). $n = 5$ biological replicates. **t**, Longevity and highest passage number of organoids from AQP5⁺ (green circles) and AQP5[−] (grey diamonds) cells. $n = 3$ biological replicates. **u**, Organoid differentiation protocol. **v–x**, Relative *Aqp5* (**v**), *Lgr5* (**w**) and *Muc5ac* (**x**) expression (by qPCR) in AQP5⁺ cell-derived organoids on day 4 of differentiation (diff). $n = 5$ biological replicates. Scale bars, 25 μ m (**b–m**), 100 μ m (**s**). Graphs represent mean \pm s.e.m. with two-sided t -test (**o**, **p**, **r**, **t**, **v–x**).

and *Aqp5-creERT2* models also faithfully reported endogenous AQP5 expression in tissues such as cornea, lung, mammary gland and salivary gland^{8,9} (data not shown).

To evaluate the utility of AQP5 as a marker for isolating enriched pyloric stem cell populations, we sorted AQP5⁺ and AQP5[−] cells from

adult wild-type mice using AQP5 antibody by fluorescence-activated cell sorting (FACS) (Fig. 2n, Extended Data Fig. 4a, b) to profile their transcriptomes and evaluate their in vitro capacity to form organoids. *Aqp5* and *Lgr5* were markedly enriched in the AQP5⁺ population relative to the AQP5[−] population, by qPCR and microarray (Fig. 2o, p,

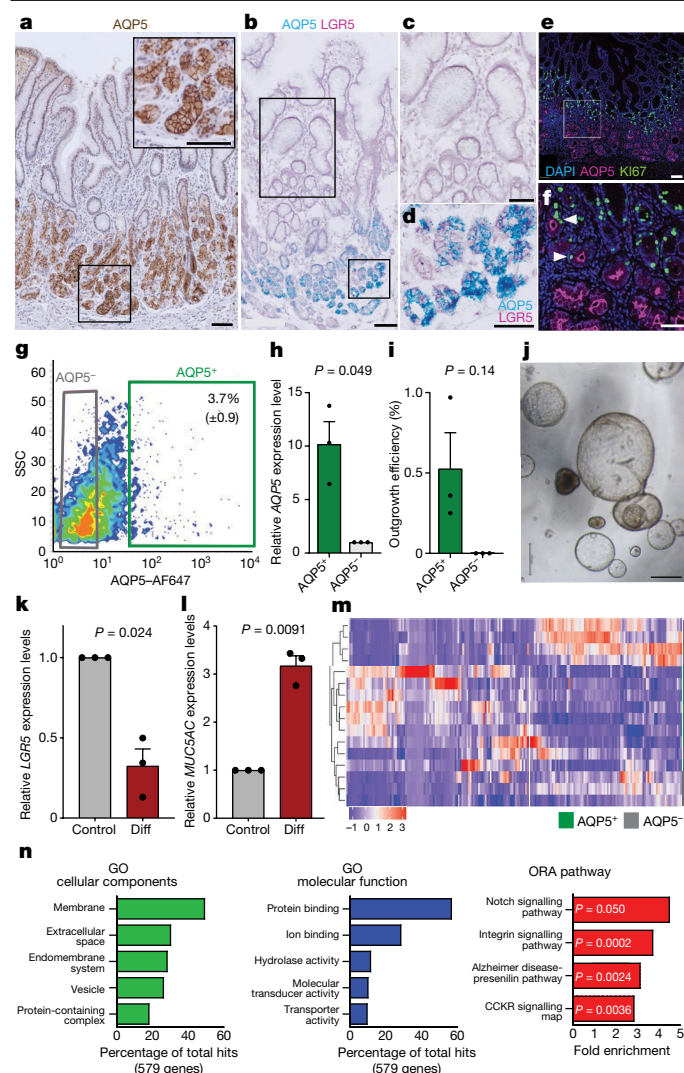


Fig. 3 | Human pyloric AQP5⁺ cells are stem cells ex vivo. **a**, Immunostaining of AQP5 in normal human pylorus. $n = 3$ biological replicates. **b–d**, AQP5 and LGR5 expression in pylorus (**b**), near gland surface (**c**) and base (**d**) by ISH. $n = 3$ biological replicates. **c** and **d** are higher-magnification images of the areas in the top and bottom black boxes, respectively, in **b**. **e, f**, AQP5 and KI67 colocalization (arrowheads). $n = 3$ biological replicates. **f** shows a higher-magnification image of the area in the white box in **e**. **g**, FACS gating for sorting human AQP5⁺ and AQP5[−] cells. Mean \pm s.e.m., $n = 4$ biological replicates. **h**, Relative AQP5 expression (by qPCR) in FACS-sorted AQP5⁺ and AQP5[−] cells. $n = 3$ biological replicates. **i, j**, Outgrowth efficiency of single AQP5⁺ and AQP5[−] cells in vitro (**i**) with a representative image (**j**). $n = 3$ biological replicates. **k, l**, Relative LGR5 (**k**) and MUC5AC (**l**) expression in (by qPCR) AQP5⁺ cell-derived organoids on day 4 of differentiation. $n = 3$ biological replicates. **m**, Heat map of 200 genes with highest variation between normal human pylorus AQP5⁺ and AQP5[−] populations by RNA-seq. $n = 8$ biological replicates. **n**, Gene Ontology (GO) terms associated with genes significantly enriched in AQP5⁺ population. ORA, overrepresentation analysis. Scale bars, 100 μ m (**a, b, e, j**), 50 μ m (**c, d, f**). Graphs are presented as mean \pm s.e.m., two-sided t -test.

Extended Data Fig. 4d, e). AQP5⁺ and LGR5-eGFP^{high} transcriptomes are highly correlated by gene set enrichment analysis (false discovery rate P value < 0.001) (Extended Data Fig. 4c), with the AQP5⁺ population presenting strong enrichment of our newly identified markers (Fig. 2q, Extended Data Fig. 4f) and the previously published gland-base markers *Lrig1*¹⁰ and *Runx1*¹¹ (Extended Data Fig. 4g). *Axin2*¹², *Cck2r* (also known as *Cck2b*)¹³ and *Sox2*¹⁴ were not enriched in the AQP5⁺ population, consistent with their relatively broad expression within pyloric

glands (Extended Data Fig. 4g). Concurring with the immunostaining data, AQP5⁺ cells presented high *Gif* (also known as *Cblif*), moderate *Ki67* (also known as *Mki67*) and no *Muc5ac* expression; there was no significant difference in *Gast* and *Chga* expression between AQP5⁺ and AQP5[−] populations, which probably reflects the limited numbers of GAST⁺ G cells and CHGA⁺ endocrine cells within the AQP5⁺ population (Extended Data Fig. 4h). Therefore, this antibody-based strategy facilitates the enrichment of mouse pyloric stem cells, independently of fluorescent reporters.

Compared to AQP5[−] cells, AQP5⁺ cells generated threefold-more organoids (0.64% versus 2.58%, respectively) that could be maintained in the long term, whereas the few organoids derived from AQP5[−] cells died within three weeks (Fig. 2r–t). Organoid initiation frequencies from AQP5⁺ cells (2.58%) and LGR5-eGFP⁺ cells (3.09%) from *Lgr5*-2A-eGFP mice were similar (Extended Data Fig. 4s–u), indicating high functional overlap. AQP5⁺ cell-derived organoids showed heterogeneous AQP5 expression, which partially overlapped with KI67 (Extended Data Fig. 4i, j)—similar to the in vivo pattern of AQP5 and KI67. Withdrawal of WNT3A, FGF10 and NOGGIN (Fig. 2u) resulted in the downregulation of the stem cell markers *Lgr5* and *Aqp5* and the upregulation of the differentiation marker *Muc5ac* in the organoids after three days (Fig. 2v–x). Therefore, AQP5 is a useful marker for the prospective isolation of enriched mouse pyloric stem cells.

AQP5 enriches for stem cells in humans

We sought to evaluate AQP5 as a marker that facilitates the enrichment of human pyloric stem cells. AQP5—along with MUC6, A4GNT and SLC9A3 (homologous to the mouse pyloric markers)—was exclusively expressed at human pyloric gland bases, overlapping with LGR5 (Fig. 3a–d, Extended Data Fig. 5a–f, a'–f'). A minor proportion of the human AQP5⁺ cells were KI67⁺ (Fig. 3e, f), reminiscent of the mouse pylorus (Extended Data Fig. 2x). Human AQP5⁺ cells overlapped with PEPC⁺ and MUC6⁺ populations, but not GIF⁺ (also known as CBLIF) parietal cells or MUC5AC⁺ foveolar cells (Extended Data Fig. 5g–j).

We next sorted AQP5⁺ cells from healthy human pyloric specimens using AQP5 antibody and verified a 10.2-fold enrichment of AQP5 expression in AQP5⁺ cells by qPCR (Fig. 3g, h). AQP5⁺ cells routinely established organoids that were passaged for more than three months, whereas AQP5[−] cells never initiated organoids (Fig. 3i, j). The human pyloric organoids expressed AQP5 heterogeneously, partially overlapping with KI67 (Extended Data Fig. 5k, l). Withdrawal of WNT3A, NOGGIN and FGF10 resulted in reduced LGR5 and AXIN2—and increased MUC5AC and TFF2—expression, indicating differentiation towards mucous lineages (Fig. 3k, l, Extended Data Fig. 5m–o).

We then performed RNA sequencing (RNA-seq) on AQP5⁺ pyloric cells isolated by FACS from healthy human pylori, and validated the top hits using qPCR. We identified more than 500 differentially expressed genes that were significantly up- or downregulated by more than fourfold in AQP5⁺ versus AQP5[−] populations (Fig. 3m, top candidates are listed in Supplementary Table 3). Enrichment of 18 candidates in the AQP5⁺ population was validated by qPCR (Extended Data Fig. 6a–e). Genes homologous to the newly identified mouse pyloric markers—including AQP5, A4GNT and MUC6 and an intestinal stem cell marker, SMOC2¹⁵—were upregulated in the AQP5⁺ fraction by RNA-seq and qPCR (Extended Data Fig. 6a). ISH confirmed SMOC2 as being expressed in a subset of gland-base cells (Extended Data Fig. 6f, f', f''). Gene Ontology analysis revealed that approximately half of the candidates were membrane-expressed and had protein-binding activity (Fig. 3n, Extended Data Fig. 6b), which suggests that additional markers could be identified from our list to potentially further enrich for pyloric stem cells. Moreover, components of key signalling pathways (for example, WNT and Notch), chemokine signalling and extracellular matrix—as well as some uncharacterized genes—were also enriched in the AQP5⁺ population, highlighting the additional biological insight to be gained from the profile (Fig. 3n, Extended Data Fig. 6c–g).

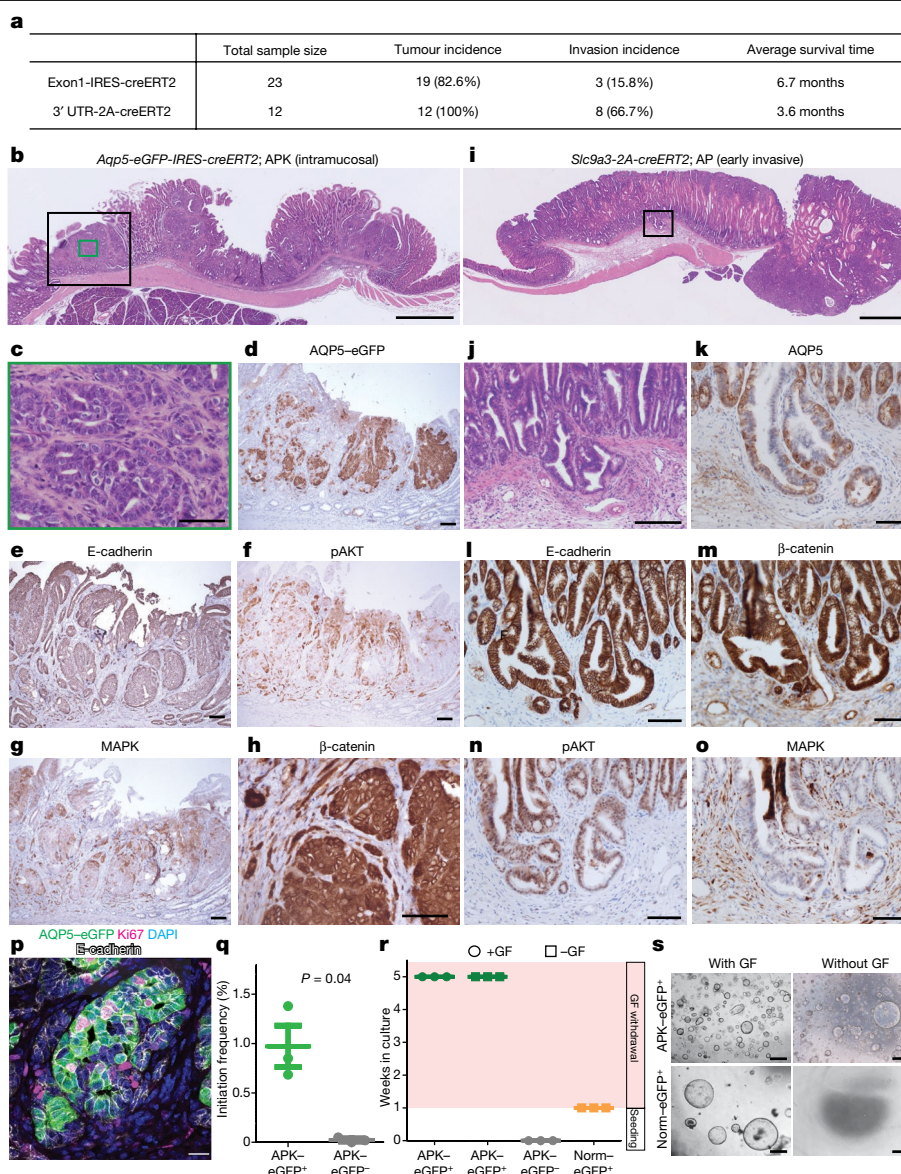


Fig. 4 | AQP5⁺ population is a source of WNT-driven, invasive gastric cancer.

a, Characteristics of models of gastric cancer. UTR, untranslated region. **b–h**, Characterization of *Aqp5-eGFP-IRES-creERT2* APK intramucosal gastric adenocarcinoma. Haematoxylin and eosin (H & E) stain of entire pyloric region (**b**) and tumour region (green box) (**c**). Immunostaining of the tumour region (black box) in **b** for AQP5–GFP (**d**), E-cadherin (**e**), phospho-(p)AKT (**f**), MAPK (**g**) and β -catenin (**h**). **h** is a higher-magnification image of Extended Data Fig. 9a. **i–o**, Characterization of *Slc9a3-2A-creERT2* AP gastric adenocarcinoma. H & E stain of the entire pyloric region (**i**) and the focal invasion (black box in **i**) (**j**). Immunostaining for AQP5 (**k**), E-cadherin (**l**), β -catenin (**m**), phospho-AKT

(**n**) and MAPK (**o**) in the focal invasion. **p**, Costaining of eGFP, K167 and E-cadherin in *Aqp5-creERT2* APK pyloric tumour. **q**, Initiation frequencies of eGFP⁺ cells (APK–eGFP⁺) versus eGFP⁺ cells (APK–eGFP[−]) from *Aqp5-creERT2* APK pyloric tumour. Paired two-sided *t*-test, mean \pm s.e.m. **r**, Mean and individual longevity of organoids derived from APK–eGFP⁺, APK–eGFP[−] and AQP5–eGFP⁺ cells from normal pylorus (norm–GFP⁺) cells seeded with exogenous growth factors (GF) for a week and without GF for the next four weeks. **s**, Representative images of organoids derived from APK–eGFP⁺ cells and norm–eGFP⁺ cells in the respective growth factor conditions. Scale bars, 1 mm (**b**, **i**), 100 μ m (**d–h**, **j–o**), 20 μ m (**p**), 500 μ m (**s**).

Collectively, these data demonstrate the utility of AQP5 as a marker for isolating enriched populations of endogenous stem cells from human stomach epithelia for downstream purposes.

AQP5⁺ cells as source of gastric cancer

We targeted conditional oncogenic mutations to pyloric-stem-cell-enriched populations using our new *creERT2* mouse models to evaluate the contribution of these populations to gastric cancer and circumvent the rapid lethality of LGR5⁺ cell-driven models of cancer. To determine how frequently pathways are co-dysregulated in gastric cancer, we analysed transcriptomes of patients with gastric cancer from The Cancer

Genome Atlas (TCGA) ($n = 155$) and East Asian cohorts ($n = 42$) for WNT/ β -catenin, PI3K and KRAS signalling activities (Extended Data Fig. 7a, b), the components of which are frequently mutated in gastric cancer^{3,16}. Using published gene signatures^{17–19}, we found that WNT/ β -catenin signalling was commonly hyperactivated in human gastric cancer (more than 80% in both cohorts of patients), and frequently co-occurred with hyperactivation of PI3K and/or RAS signalling (57.1–64.3%) (Extended Data Fig. 7). We thus recapitulated these co-dysregulated pathways by crossing our pyloric *creERT2* drivers to conditional *Apc*, *Pten* and *Kras*^{G12D} alleles, and induced hyperactivation of the pathways at adulthood.

All mouse models developed sizeable tumours exclusively in the pylorus, with latencies ranging from 1 to 11 months after induction (in

$n = 30$ out of 34 mice) (Extended Data Fig. 8a–g). In the tumours, which were classified as tubular-type gastric adenocarcinomas (according to the World Health Organization (WHO) classification), malignant structures surrounded by stroma and inflammatory cells replaced normal glands. Sole hyperactivation of *Kras*^{G12D} did not produce pyloric tumours (Extended Data Fig. 8h, i). The pylori of uninduced mice of the same cancer genotype that lacked Cre were normal (Extended Data Fig. 8j).

Across all combinations of oncogenes, the tumour incidence in 2A-creERT2 models was 100% compared to 82.6% in IRES-creERT2 models (Fig. 4a), reflecting differences in Cre activation efficacies. The 2A-creERT2 models displayed almost contiguous tumour growth, which contrasts with the multifocal lesions in IRES-creERT2 models (Extended Data Fig. 8b–g). Hyperactivation of WNT/ β -catenin signalling alone (*Apc*^{R/R} (hereafter, A)) was sufficient to drive tumorigenesis, and co-activation of the PI3K and/or KRAS pathways (*Apc*^{R/R}; *Pten*^{R/R} (hereafter AP) or *Apc*^{R/R}; *Pten*^{R/R}; *Kras*^{G12D} (hereafter, APK)) accelerated tumour development and progression (Extended Data Fig. 8a–g). We also observed focal invasions through the muscularis mucosae in A and AP models (IRES-creERT2, 15.8% and 2A-creERT2, 66.7%) (Fig. 4a). As expected, intestinal tumours were never observed (Extended Data Fig. 8k, l).

We characterized all tumours to detail pathway activation, proliferation status, lineage-marker expression and epithelia and stroma constitution. As there are no major phenotypic differences between the models, we present data from *Aqp5*-IRES-creERT2 APK tumours and focal invasions from *Slc9a3*-2A-creERT2 AP tumours (Fig. 4b, c, i, j). The pyloric tumours and focal invasions were predominantly tdTomato⁺, confirming that they originated from AQP5⁺ or SLC9A3⁺ cells (Extended Data Fig. 9b, j). In contrast to the normal pylorus (Extended Data Fig. 9p–w), the gastric adenocarcinomas presented hyperactivation of WNT/ β -catenin, MAP kinase (MAPK) and phospho-AKT pathways, as evidenced by increased levels of expression of nuclear and/or cytoplasmic β -catenin, MAPK and phospho-AKT, respectively (Fig. 4f–h, m–o, Extended Data Fig. 9a). These regions were also highly proliferative, and lacked GIF, GAST and MUC5AC expression (Extended Data Fig. 9d–g, l–o). In the *Aqp5*-IRES-creERT2 APK model, tdTomato⁺ cells that retained E-cadherin expression were also found throughout the tumour stroma (Fig. 4e). In the *Slc9a3*-2A-creERT2 AP model, E-cadherin⁺ cells infiltrated through the muscularis mucosae (Fig. 4l). Immunostaining for eGFP reporter expression revealed *Aqp5* expression in a subpopulation of the pyloric tumours, some of which were Ki67⁺ (Fig. 4d, k, p). Many of the AQP5⁺ cells in the tumours co-expressed *Lgr5*, with increased *Aqp5* expression in tumours compared to adjacent normal mucosa (Extended Data Fig. 9h, h', h''). There was a low incidence of tumours within non-gastrointestinal organs, such as the salivary gland (less than 25% in *Aqp5*-IRES-creERT2-driven cancer models), which did not affect survival to preclude the development of gastric adenocarcinoma (Extended Data Fig. 9i, i'). Thus, our mouse models support pyloric-stem-cell-enriched populations as being a source of invasive, WNT-driven gastric cancer and are valuable for modelling gastric cancer in vivo.

AQP5⁺ tumour cells show ex vivo stemness

We then determined whether AQP5⁺ tumour cells behave differently from their AQP5[−] counterparts. After confirming that the eGFP reporter recapitulated endogenous AQP5 expression in the *Aqp5*-IRES-creERT2 APK tumours (Extended Data Fig. 9x, x'), we cultured sorted eGFP⁺ and eGFP[−] cells (Extended Data Fig. 9y, z, z', z''). eGFP⁺ tumour cells reproducibly generated organoids that could be serially propagated in the absence of exogenous growth factors, whereas AQP5[−]-GFP[−] tumour populations never produced organoids, despite containing Ki67⁺ cells (Fig. 4p–s). Moreover, although AQP5[−]-GFP⁺ normal cells could initiate organoids with growth factors, they died upon the removal of growth factors (Fig. 4r, s). These data suggest that the stem potential of the tumour is found within the AQP5⁺ compartment.

AQP5 expression in human gastric cancer

We surveyed AQP5 expression in human gastric cancer by immunostaining on a tissue microarray of 145 samples of distal gastric cancer comprising intestinal, diffuse and mixed subtypes with variable grades of differentiation. AQP5 was expressed in most intestinal, diffuse and mixed cases (Extended Data Fig. 10a–e). Contrasting with normal pylorus, 96.1% of the tumour samples displayed cytoplasmic AQP5 expression, and 37.9%, 3.9% and 37.9% had membranous, nuclear and multiple localizations of AQP5, respectively (Extended Data Fig. 10a–e). Although the intracellular localization of AQP5 has previously been reported in other cancers^{20–22}, its functional relevance is unknown.

Full sections of 54 advanced human distal gastric adenocarcinomas and 12 metastatic lesions showed that most expressed AQP5 (Extended Data Fig. 10f). Cytoplasmic AQP5 was observed in all AQP5⁺ samples, and membranous or luminal, nuclear and multiple sites of AQP5 localization were found in 53.7%, 9.8% and 53.7% of the samples, respectively (Extended Data Fig. 10f–j). In 51.2% of the sections, submucosal malignant cells expressed more AQP5 than their mucosal counterparts (Extended Data Fig. 10g–j, o, p). AQP5 was expressed in poorly cohesive tumour cells in 70% of the cohort (Extended Data Fig. 10j, o). All AQP5⁺ cells in the submucosa retained E-cadherin expression, and a subset co-expressed Ki67 (Extended Data Fig. 10k–n). In cases of intestinal metaplasia (which is strongly correlated with gastric cancer^{23,24}), 46.2% displayed mild-to-moderate AQP5 expression (Extended Data Fig. 10o, q). AQP5 was also weakly expressed in a subset of signet ring cells in 66.7% of the samples (Extended Data Fig. 10o, r). Of the metastatic lesions with AQP5⁺ primary tumours, 83.3% contained AQP5⁺ tumour cells in the lymph node (Extended Data Fig. 10o, s).

Our broad survey shows that AQP5 is commonly expressed in primary intestinal and diffuse subtypes of gastric cancer, as well as in metastases of these subtypes.

Efforts to exploit the therapeutic potential of stem cells require functionally validated markers for their prospective isolation. To our knowledge, for the first time we have purified enriched populations of human pyloric stem cells, and demonstrated the direct contribution of AQP5⁺ cells in mouse models of gastric carcinogenesis. Although the expression of secretory markers such as GIF and MUC6 may not conform to the classical 'undifferentiated' stem cell model, it is well-established in the liver and lung that more-specialized cells can serve as homeostatic stem cells^{25–28}, an emerging trend in epithelial organs²⁹. AQP5 is known for regulating water transport in healthy tissues⁶, and has increasingly been implicated in cancers as a driver of proliferation and invasiveness in vitro^{30–34}. Various human cancers—including gastric, breast, soft tissue sarcoma, lung, oesophageal and colorectal cancers—present high AQP5 expression^{21,22,33,35–37}. We show that AQP5 is expressed in most human primary tumours, and metastases, of intestinal and diffuse subtypes of gastric cancer. We found that tumour-resident AQP5⁺ cells in our mouse model of gastric cancer selectively exhibited ex vivo stem potential, indicating that the AQP5⁺ tumour population contains cancer stem cells. Future evaluation of the stem potential of AQP5⁺ in human gastric cancers using our antibody-based isolation protocols has the potential to reveal opportunities for developing more-effective cancer therapeutic strategies.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-1973-x>.

1. Barker, N. et al. *Lgr5*^{Cre} stem cells drive self-renewal in the stomach and build long-lived gastric units in vitro. *Cell Stem Cell* **6**, 25–36 (2010).

2. Barker, N. et al. Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* **457**, 608–611 (2009).
3. The Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* **513**, 202–209 (2014).
4. Barker, N. et al. Identification of stem cells in small intestine and colon by marker gene *Lgr5*. *Nature* **449**, 1003–1007 (2007).
5. Tian, H. et al. A reserve stem cell population in small intestine renders *Lgr5*-positive cells dispensable. *Nature* **478**, 255–259 (2011).
6. Direito, I., Madeira, A., Brito, M. A. & Soveral, G. Aquaporin-5: from structure to function and dysfunction in cancer. *Cell. Mol. Life Sci.* **73**, 1623–1640 (2016).
7. Grün, D. et al. Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* **525**, 251–255 (2015).
8. Funaki, H. et al. Localization and expression of AQP5 in cornea, serous salivary glands, and pulmonary epithelial cells. *Am. J. Physiol.* **275**, C1151–C1157 (1998).
9. Matsuzaki, T. et al. Expression and immunolocalization of water-channel aquaporins in the rat and mouse mammary gland. *Histochem. Cell Biol.* **123**, 501–512 (2005).
10. Choi, E. et al. Lrig1+ gastric isthmal progenitor cells restore normal gastric lineage cells during damage recovery in adult mouse stomach. *Gut* **67**, 1595–1605 (2018).
11. Matsuo, J. et al. Identification of stem cells in the epithelium of the stomach corpus and antrum of mice. *Gastroenterology* **152**, 218–231 (2017).
12. Sigal, M. et al. Stromal R-spondin orchestrates gastric epithelial stem cells and gland homeostasis. *Nature* **548**, 451–455 (2017).
13. Hayakawa, Y. et al. CCK2R identifies and regulates gastric antral stem cell states and carcinogenesis. *Gut* **64**, 544–553 (2015).
14. Arnold, K. et al. Sox2+ adult stem and progenitor cells are important for tissue regeneration and survival of mice. *Cell Stem Cell* **9**, 317–329 (2011).
15. Muñoz, J. et al. The *Lgr5* intestinal stem cell signature: robust expression of proposed quiescent +4' cell markers. *EMBO J.* **31**, 3079–3091 (2012).
16. Muratani, M. et al. Nanoscale chromatin profiling of gastric adenocarcinoma reveals cancer-associated cryptic promoters and somatically acquired regulatory elements. *Nat. Commun.* **5**, 4361 (2014).
17. Zhang, Y. et al. A pan-cancer proteogenomic atlas of PI3K/AKT/mTOR pathway alterations. *Cancer Cell* **31**, 820–832 (2017).
18. Pek, M. et al. Oncogenic KRAS-associated gene signature defines co-targeting of CDK4/6 and MEK as a viable therapeutic strategy in colorectal cancer. *Oncogene* **36**, 4975–4986 (2017).
19. Van der Flier, L. G. et al. The intestinal Wnt/TCF signature. *Gastroenterology* **132**, 628–632 (2007).
20. Lambert, N. et al. Expression of aquaporin 5 and the AQP5 polymorphism A(-1364)C in association with peritumoral brain edema in meningioma patients. *J. Neurooncol.* **112**, 297–305 (2013).
21. Shimizu, H. et al. The expression and role of aquaporin 5 in esophageal squamous cell carcinoma. *J. Gastroenterol.* **49**, 655–666 (2014).
22. Zhu, Z. et al. Expression of AQP3 and AQP5 as a prognostic marker in triple-negative breast cancer. *Oncol. Lett.* **16**, 2661–2667 (2018).
23. Kinoshita, H., Hayakawa, Y. & Koike, K. Metaplasia in the stomach—precursor of gastric cancer? *Int. J. Mol. Sci.* **18**, 2063 (2017).
24. Jencks, D. S. et al. Overview of current concepts in gastric intestinal metaplasia and gastric cancer. *Gastroenterol. Hepatol.* **14**, 92–101 (2018).
25. Rawlins, E. L., Clark, C. P., Xue, Y. & Hogan, B. L. M. The *Id2*+ distal tip lung epithelium contains individual multipotent embryonic progenitor cells. *Development* **136**, 3741–3745 (2009).
26. Barkauskas, C. E. et al. Type 2 alveolar cells are stem cells in adult lung. *J. Clin. Invest.* **123**, 3025–3036 (2013).
27. Desai, T. J., Brownfield, D. G. & Krasnow, M. A. Alveolar progenitor and stem cells in lung development, renewal and cancer. *Nature* **507**, 190–194 (2014).
28. Wang, B., Zhao, L., Fish, M., Logan, C. Y. & Nusse, R. Self-renewing *Axin2*+ cells fuel homeostatic renewal of the liver. *Nature* **524**, 180–185 (2015).
29. Post, Y. & Clevers, H. Defining adult stem cell function at its simplest: the ability to replace lost cells through mitosis. *Cell Stem Cell* **25**, 174–183 (2019).
30. Kang, S. K. et al. Aquaporin 5 (AQP5) is a novel signaling molecule triggering Ras/ERK/retinoblastoma (Rb) signaling pathway in colon cancer cell lines. *Mol. Cancer Ther.* **6**, 8115 (2007).
31. Woo, J. et al. The effect of aquaporin 5 overexpression on the Ras signaling pathway. *Biochem. Biophys. Res. Commun.* **367**, 291–298 (2008).
32. Woo, J. et al. Overexpression of AQP5, a putative oncogene, promotes cell growth and transformation. *Cancer Lett.* **264**, 54–62 (2008).
33. Chae, Y. K. et al. Expression of aquaporin 5 (AQP5) promotes tumor invasion in human non small cell lung cancer. *PLoS ONE* **3**, e2162 (2008).
34. Huang, Y.-H. et al. Aquaporin 5 promotes the proliferation and migration of human gastric carcinoma cells. *Tumour Biol.* **34**, 1743–1751 (2013).
35. Shimasaki, M., Kanazawa, Y., Sato, K., Tsuchiya, H. & Ueda, Y. Aquaporin-1 and -5 are involved in the invasion and proliferation of soft tissue sarcomas. *Pathol. Res. Pract.* **214**, 80–88 (2018).
36. Wang, W. et al. Expression of AQP5 and AQP8 in human colorectal carcinoma and their clinical significance. *World J. Surg. Oncol.* **10**, 242 (2012).
37. Shen, L. et al. Expression profile of multiple aquaporins in human gastric carcinoma and its clinical significance. *Biomed. Pharmacother.* **64**, 313–318 (2010).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Article

Methods

Mice

For exon 1 knock-ins, an eGFP-IRES-creERT2 cassette was inserted immediately downstream of the start codons of *Aqp5*, *A4gnt*, *Spp1* and *Slc9a3* gene loci by homologous recombination in embryonic stem cells, as illustrated in Extended Data Fig. 2a. For 3' UTR knock-ins, a 2A-creERT2, 2A-eGFP or 2A-DTR cassette was inserted immediately before the stop codon of *Aqp5*, *Lgr5* and/or *Slc9a3* gene loci by homologous recombination in embryonic stem cells, as illustrated in Extended Data Fig. 2b, thereby preserving the intact protein-coding region and endogenous expression of the genes. *Rosa26-tdTomato^{LSL}* (Ai14) (JAX ID 007914) mice³⁸ were obtained from Jackson Laboratories. *Lgr5-eGFP-IRES-creERT2* (ref. 4) (JAX ID 008875), *Lgr5-DTR-eGFP* (MGI ID 5294798)⁵, *Kras^{LSL-G12D}* (JAX ID 019104)³⁹, *Apc^{fl/fl}* (MGI ID 1857966)⁴⁰ and *Pten^{fl/fl}* (MGI ID 2182005)⁴¹ mice have previously been described. All Cre and eGFP lines were bred as heterozygotes except *A4gnt-IRES-creERT2* and *Spp1-IRES-creERT2* mice, which were bred as homozygotes. All mouse experiments were approved by the Institutional Animal Care and Use Committee of A*STAR, and performed in compliance with all relevant ethical regulations. The maximum tumour size allowed by the IACUC is 20 mm in any dimension and none of the experiments exceeded this limit. For all experiments, adult mice (not selected for sex) with a minimum age of 7–8 weeks were used. The experiments were not randomized, and there was no blinded allocation during experiments and outcome assessment. No statistical methods were used to predetermine sample size. Genotyping primers are collated in Supplementary Table 5. Mouse lines are available from N.B. upon request.

Human material

Normal human pylorus for FACS was provided by K. G. Yeoh, J. So and A. Shabbir, NUS Department of Medicine and Pathology (granted under IRB protocol 11-167E) and N. Inaki and T. Tsuji, Ishikawa Prefectural Central Hospital. Informed consent was obtained from all patients and experiments were performed in compliance with all relevant ethical regulations. Human distal cancer formalin-fixed paraffin-embedded (FFPE) sections were provided by NUS Department of Medicine and Pathology (granted under IRB protocol 11-167E) and Leeds Teaching Hospitals NHS Trust (granted under IRB protocol CA01_122).

Mouse treatment

Mice were each injected with tamoxifen dissolved in sunflower oil intraperitoneally, at 4 mg tamoxifen per 30 g body weight. Diphtheria-toxin-treated mice were injected with a single dose of diphtheria toxin dissolved in PBS intraperitoneally, at 0.5 µg diphtheria toxin per 30 g body weight.

Gland isolation, cell dissociation and flow cytometry

Mouse pylorus. Mouse pylorus was incubated in chelation buffer (5.6 mM sodium phosphate, 8 mM potassium phosphate, 96.2 mM sodium chloride, 1.6 mM potassium chloride, 43.4 mM sucrose, 54.9 mM D-sorbitol, 1 mM dithiothreitol) with 5 mM EDTA at 4 °C for 2 h. Glands were isolated by repeated pipetting of finely chopped pylorus tissue in cold chelation buffer. Chelation buffer containing isolated glands was filtered through a 100-µm filter mesh, and centrifuged at 720g at 4 °C for 3 min. The pellet was resuspended in TrypLE (Life Technologies) with DNaseI (0.8 U/µl) (Sigma) and incubated at 37 °C for 10 min with intermittent trituration for digestion into single cells. Digestion was quenched by dilution with cold HBSS buffer. The suspension was centrifuged at 720g at 4 °C for 3 min. For AQP5 antibody stain, the pellet was resuspended in HBSS with 2% fetal bovine serum (FBS) (Hyclone) with AQP5-AF647 (Abcam, ab215225) at 1:500 dilution and incubated on ice at 30 min in the dark. The pellet was subsequently washed twice with cold HBSS and spun at 800g for 3 min at 4 °C. The

pellet was resuspended in HBSS with 2% FBS. Before sorting, 1 µg/ml propidium iodide (Life Technologies) was added to the cell suspensions, filtered through a 40-µm strainer and sorted on BD Influx Cell Sorter (BD Biosciences). Cells were collected in RLT Plus buffer (Qiagen) for RNA extraction or HBSS with 2% FBS and 1% PenStrep (Gibco) for organoid culture.

Human pylorus. Human pylorus was collected in advanced DMEM/F-12 medium with 10 mM HEPES, 2 mM Glutamax (incubation buffer, all from Life Technologies), supplemented with 1× Anti-Anti (Life Technologies) and 1 mM *N*-acetylcysteine (Sigma). After at least 3 washes in HBSS, the pylorus was finely chopped and digested in incubation buffer supplemented with 1 mg/ml collagenase (Gibco) and 2 mg/ml bovine serum albumin (Sigma) for 30 min at 37 °C with intermittent mixing. The remainder of the processing protocol is identical to that for the mouse tissue described in 'Mouse pylorus'. Cells for organoid culture were collected in organoid culture medium with growth factors and 0.2% growth-factor-reduced Matrigel (Corning) (v/v).

Organoid culture

Organoid culture of FACS-isolated single human and mouse pylorus cells were performed as previously described⁴². In brief, single cells were resuspended in growth-factor-reduced Matrigel (Corning) and cultured in basal medium (advanced DMEM/F-12 medium with 10 mM HEPES, 2 mM Glutamax, 1× N2, 1× B27 (all Invitrogen), *N*-acetyl-cysteine (Sigma) and primocin (Invivogen)) supplemented with the following growth factors: EGF (Invitrogen), GAST (Sigma), FGF10 (Peprotech), Noggin (Peprotech), WNT3A (Millipore), R-spondin and ROCK inhibitor Y27632 (Sigma). A83-01 (Tocris) was also added to human pyloric cultures. Mouse cancer organoids were grown in only basal medium after first week of culture. Organoids were passaged when confluent, at least once a week. Only organoids beyond 100 µm and 200 µm in diameter with a clear central lumen are scored as organoids for Fig. 2 and Fig. 3, respectively.

RNA isolation and qPCR

Tissues were lysed in Trizol (Qiagen) and single cells were lysed in RLT Plus buffer (Qiagen). RNA was subsequently isolated with RNeasy Universal Plus kit (Qiagen) and cDNA was generated with Superscript III (Life Technologies) according to the manufacturer's instructions. qPCR was performed with a minimum of three biological replicates per gene using SYBR green dye (Promega) according to the manufacturer's instructions, and ran on StepOne or Quantstudio7 qPCR machines (Applied Biosystems). Analysis was carried out using the double *C_t* method on Step One Software on the respective qPCR machines (Applied Biosystems). qPCR validation of top candidates from RNA-seq was performed on 2–4 ng of SPIA-amplified cDNA derived from the Ovation Pico WTA system (Nugen Technologies) owing to limitation of RNA availability. In the event that any of the samples for a specific target does not amplify, the relative expression values of all the samples for that target are increased by one to enable visualization of the values on a log scale. Sequences of qPCR primers are collated in Supplementary Table 5.

Transcriptome profiling and analysis

Single-cell RNA-seq, CEL-seq and RaceID. Single LGR5-eGFP^{high} pyloric epithelial cells from *Lgr5-DTR-eGFP* mice⁵ were isolated by FACS (as described in 'Mouse pylorus' in 'Gland isolation, cell dissociation and flow cytometry') and collected in each well of 96-well plates. Total RNA extracted from each cell was used to generate single-cell RNA expression libraries as previously described⁷. A total of 285 LGR5-eGFP^{high} cells from 3 mice were sequenced on Illumina HighSeq 2500 instrument using 101 base-pair paired-end sequencing. *K*-means clustering in RaceID was used to delineate clusters of subpopulations, as previously described⁷.

Microarray and analysis. Labelling, hybridization and washing protocols for microarrays were performed according to Origene instructions. RNA quality was first determined by assessing the integrity of the 28S and 18S ribosomal RNA bands on Agilent RNA 60000 Pico LabChips in an Agilent 2100 Bioanalyzer (Agilent Technologies). A minimum of 2 ng of RNA was used to generate SPIA-amplified cDNA using the Ovation Pico WTA system (Nugen Technologies). Five micrograms of SPIA-amplified purified cDNA was then fragmented and biotin-labelled using the Nugen Encore Biotin module (Nugen Technologies). Microarray was performed using the Affymetrix Mouse ST v.2.0 GeneChips (Affymetrix), which consists of more than 28,000 probes for previously annotated genes. The individual microarrays were washed and stained in an Affymetrix Fluidics Station 450, and hybridized probe fluorescence was detected using the Affymetrix G3000 GeneArray Scanner. Image analysis was carried out on the Affymetrix GeneChip Command Console v.2.0 using the MAS5 algorithm. CEL files were generated for each array and used for gene-expression analysis. The CEL files were then processed in R (v.3.2.3) with the Bioconductor (v.3.2) libraries 'oligo' (v.1.34.2), 'pd.mogene.2.0.st' (v.3.14.1) and 'limma' (v.3.26.8). We used robust multi-array average to perform background correction and normalization with the 'rma' function implemented in the 'oligo' package ('target' parameter was set to 'core' to obtain expression values at the gene level). The experimental design was stored as a single factor with individual levels for each combination of LGR5-GFP level (high, low or negative) or AQP5 status (positive or negative). Linear models were fitted to the expression data with the function 'lmFit' (default parameters). The relevant contrasts were fitted with 'contrasts.fit' (default parameters); differential expression was tested with 'eBayes' (default parameters). Differential gene expression was analysed using Partek Genomics Suite software (Partek). Relative gene expressions are depicted as single values as given by Partek analysis software. Gene set enrichment analysis was performed using the GSEA v.6.1^{43,44}.

Rna-seq. AQP5⁺ and AQP5⁻ cells were collected directly into RLT Plus buffer by FACS sorting. Total RNA was isolated using Qiagen RNeasy MicroKit (Qiagen). RNA quality was first determined by assessing the integrity of the 28S and 18S ribosomal RNA bands on Agilent RNA 60000 Pico LabChips in an Agilent 2100 Bioanalyzer (Agilent Technologies). Amplified cDNA library was prepared according to the manufacturer's instructions with SMARTer Stranded Total RNA-Seq Kit v.2 - Pico Input Mammalian (Takara) using 10 ng of input total RNA. Indexed 150-bp paired-end sequencing was performed on HiSeq 2500 (Illumina) and Illumina real-time analysis software was used for base-calling to generate FASTQ files. The reads were mapped to Genome Reference Consortium Human Build 38 patch release 12 (GRCh38.p12) with STAR software version 2.5.3a with the following options issued: -outFilterType BySJout, -outFilterMultimapNmax 10, -alignSJoverhangMin 15, -alignSJDBoverhangMin 1, -outFilterMismatchNmax 12, -outFilterMatchNminOverLread 0.4, -alignIntronMin 20, -alignIntronMax 2000000, -outSAMattrIHstart 0, -outSAMmapqUnique 244, -outMultimapperOrder Random, -outReadsUnmapped None, -outFilterIntronMotifs None, -outSAMmode Full, -outSAMattributes All -quantMode GeneCounts, -clip3pAdapterSeq AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT. Counts per sample were subsequently concatenated in statistical software, R version 3.2.3, and reads were normalized with trimmed mean of *M* values normalization as implemented in edgeR version 3.12.1 (with limma 3.26.9). Differential expression testing was performed with the edgeR function 'glmQLFit' using a design matrix that took sample batches and AQP5 status into account. Differentially expressed genes were those with more than twofold change between AQP5⁺ and AQP5⁻ samples, with false discovery rate < 0.05. Owing to the likely inclusion of immune cells in the profile, immune-related genes⁴⁵ were omitted, resulting in a final list of >500 differentially expressed genes. Gene Ontology,

overrepresentation analysis and PANTHER pathway analysis of the differentially expressed genes were performed on the PANTHER classification system⁴⁶ using default parameters.

Transcriptomic pathway signature analysis of human gastric cancer level 3 TCGA RNA-seq normalized matrix for 415 GC and 35 normal gastric samples, and their corresponding clinical information, were downloaded from the Broad Institute TCGA Genome Data Analysis Center Firehose (<https://gdac.broadinstitute.org/>). Gene expression data of 200 GC and 100 matched normal gastric samples were generated using Affymetrix Human Genome U133 Plus 2.0 Array (GSE15459) and processed as previously described¹⁶. All normal samples, and only tumours of antral or pyloric origin, were included for analysis. To determine activity of PI3K, WNT and KRAS pathways in primary tumours, we used published pathway signatures by several groups: KRAS signature based on differential gene-expression analysis between colorectal cancers with high KRAS mutation, and wild-type KRAS tumours¹⁸; PI3K signature composed of genes modulated in vitro by PI3K inhibitors, according to the CMap signature¹⁷; and finally, intestinal WNT signature defined by profiling colorectal cancer cell lines carrying an inducible block of the WNT pathway and differential gene-expression analysis of human colon adenoma and adenocarcinomas versus normal colonic epithelium¹⁹. For each pathway signature, only upregulated genes in pathway activation were selected for downstream analysis.

To quantify the relative activation level in a specific oncogenic pathway, we derived a 'μ score' for each sample profile. In brief, the transcriptomic μ score was defined as the average of standardized expression values of those genes upregulated in a specific oncogenic pathway (after the log-transformed values centred to the standard deviation from the median across the samples included in the analysis). For each oncogenic pathway, μ scores were calculated for all normal samples, and the μ score at the 90% percentile of the normal samples was used as the cut-off to define a pathway as hyperactivated. The μ score for each tumour was determined and a μ score higher than the cut-off was considered to be hyperactivated for that particular pathway. For each combination of pathways (WNT, WNT and PI3K or WNT, PI3K and KRAS), the concurrence rate was given by the frequency of tumours that were hyperactivated in all the pathways in question.

Histology

Immunohistochemistry and immunofluorescence. Immunohistochemistry (IHC) and immunofluorescence were performed according to standard protocols. In summary, tissues were fixed in 4% paraformaldehyde in PBS (w/v) overnight at 4 °C, and processed into paraffin blocks. Eight-micrometre sections from the paraffin blocks and tissue microarray slides were deparaffinated and rehydrated, followed by antigen retrieval via heating to 121 °C in a pressure cooker in standard 10 mM citric acid pH6 buffer, a commercial citrate pH 6.1 buffer (S1699, DAKO) or Tris/EDTA buffer, pH 9.0 (S2367, DAKO). Primary antibodies used were chicken anti-EGFP (1:2,000, Abcam, ab290), rabbit anti-EGFP (1:200; Cell Signalling, 2956S), rabbit anti-KI67 (1:200; Thermofisher, MA5-14520), rabbit anti-GIF (1:10,000; provided by D. H. Alpers), rabbit anti-RFP (1:200; Rockland, 600-401-379), rabbit anti-aquaporin 5 (1:200, Santa Cruz, SC-28628 and 1:500, LSBio, LS-C756566), rabbit anti-SLC9A3 (1:200, Santa Cruz, SC-16103-R), rabbit anti-mucin 6 (1:200, LSBio, LS-C312108), rabbit anti-A4GNT (1:500, Novus Biologicals, NBPI-89129), rabbit anti-GAST (1:200, Leica Biosystems, NCL-GASp), mouse anti-MUC5AC (1:200, Leica Biosystems, NCL-HGM-45-M1), rabbit anti-vimentin (1:500, Abcam, ab92547), mouse anti-E-cadherin (1:200, BD Transduction Laboratories, 610181), mouse anti-β-catenin (1:200, BD Transduction Laboratories, 610154), mouse anti-RFP (1:200, Abcam, 129244), mouse anti-CHGA (1:200, Abcam, 15160), rabbit anti-phospho-MAPK (1:200, Cell Signalling, 4370S), mouse anti-H-K-ATPase (1:1,000, MBL International, D032-3) and rabbit anti-phospho-AKT (1:200, Cell Signalling, 3787L). Detailed information about clone number, and antibody validation can be found in Supplementary Table 6. The peroxidase-conjugated

Article

secondary antibodies used were mouse or rabbit EnVision+ (DAKO) for HRP IHC, or anti-chicken, -rabbit or -mouse Alexa 488, 568 or 647 IgG (1:500, Invitrogen) for immunofluorescence. GSII-lectin-AF568 (1:500, ThermoFisher) was incubated on the slides for 1 h at room temperature together with secondary antibodies. IHC sections were dehydrated, cleared and mounted with DPX (Sigma) and immunofluorescence sections were mounted in Hydromount (National Diagnostics) with Hoechst for nuclear staining. Immunostainings and imaging were performed on a minimum of three biological replicates and representative images of the replicates are included in the manuscript.

H&E. H & E staining was performed on FFPE sections according to standard laboratory protocols.

Whole mount analysis and vibratome sectioning. Tissues were fixed in 4% paraformaldehyde in PBS (w/v) overnight at 4 °C. Whole-mount tissues were permeabilized in 2% TritonX-100 in PBS (v/v) overnight at 4 °C, and 500-µm vibratome sections were generated by sectioning tissues embedded in 4% low-melting point agarose with a vibrating microtome (Leica) and permeabilized in 2% TritonX-100 in PBS (v/v) overnight at 4 °C. Rapiclear (Sunjin laboratory) was used to clear whole-mount tissues and vibratome sections according to the manufacturer's instructions. Hoescht was used as a nuclear counterstain.

ISH. ISH and co-ISH were performed using RNAscope⁴⁷ 2.5 High Definition Brown Assay and 2.5 High Definition Duplex Reagent Assay (Advanced Cell Diagnostics), respectively, according to the manufacturer's instructions. DapB was used as negative control for all the RNAscope experiments. ISH and imaging were performed on a minimum of three biological replicates and representative images of the replicates are included in the manuscript.

Analysis and scoring of staining on mouse and human FFPE sections. Overlap of SLC9A3-eGFP and AQP5-antibody stains was determined by counting stained cells contacting the lumen at the base of glands. The entire height of the gland base surrounding the gland base lumen had to be visible to avoid over- or underrepresentation of localized populations.

Samples of H & E sections of mouse gastric tumours were evaluated by qualified veterinary and clinical histopathologists. Scoring of AQP5 staining on human gastric cancer specimens was performed by qualified clinical histopathologists. The tumour in the tissue section was considered positive for AQP5 if staining was observed in more than 5% of the malignant cells. Subcellular localization of AQP5, relative staining intensities and stained features were all determined by qualified pathologists.

Microscopy imaging

Image acquisition. IHC and H & E slides were imaged with Zeiss AxioImager Z1 Upright microscope. RNAscope slides and large-area images were captured with Nikon Ni-E microscope and DS-Ri2 camera. Immunofluorescence slides were imaged using Olympus FV1000 and FV3000 confocal microscopes. Cultured organoids were imaged with Olympus DP-27 camera on Olympus IX53 inverted microscope.

Image processing. RNAscope and images with large areas were processed with NIS-Elements AR software (Nikon) with EDF and stitching features, respectively. Immunofluorescence images were processed using ImageJ (NIH) and whole-mount organoid images were processed using Imaris 8.0 (Bitplane).

Statistics and reproducibility

Gene-expression data were quantified and depicted as mean ± s.e.m. Statistical analyses were performed using GraphPad Prism. Data were tested for statistical significance by paired two-tailed *t*-test, unless otherwise stated in figure legends. Statistical significance of overlap between the two gene sets in Fig. 1b was determined by hypergeometric

distribution (http://nemates.org/MA/progs/overlap_stats.html). Precise *P* values of statistical significance are shown in the respective figures. Representative images of all histological experiments and FACS strategies were performed at least thrice independently, with similar results,

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Microarray data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) under accession code GSE121803. RNA-seq data of AQP5⁺ and AQP5⁻ human samples have also been deposited in the GEO, under accession code GSE133036. Source Data for Figs. 1–4 and Extended Data Figs. 1, 2, 4–6, 9 are provided with the paper. Any other relevant data supporting the findings of this study are available from the corresponding author on reasonable request.

38. Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* **13**, 133–140 (2010).
39. Jackson, E. L. et al. Analysis of lung tumor initiation and progression using conditional expression of oncogenic *K-ras*. *Genes Dev.* **15**, 3243–3248 (2001).
40. Shibata, H. et al. Rapid colorectal adenoma formation initiated by conditional targeting of the *Apc* gene. *Science* **278**, 120–123 (1997).
41. Suzuki, A. et al. T cell-specific loss of Pten leads to defects in central and peripheral tolerance. *Immunity* **14**, 523–534 (2001).
42. Leushacke, M. et al. *Lgr5*-expressing chief cells drive epithelial regeneration and cancer in the oxyntic stomach. *Nat. Cell Biol.* **19**, 774–786 (2017).
43. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
44. Mootha, V. K. et al. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
45. Monaco, G. et al. RNA-seq signatures normalized by mRNA abundance allow absolute deconvolution of human immune cell types. *Cell Rep.* **26**, 1627–1640.e7 (2019).
46. Mi, H., Muruganujan, A., Casagrande, J. T. & Thomas, P. D. Large-scale gene function analysis with the PANTHER classification system. *Nat. Protocols* **8**, 1551–1566 (2013).
47. Wang, F. et al. RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J. Mol. Diagn.* **14**, 22–29 (2012).

Acknowledgements The authors thank staff at the IMB-IMU and the SBIC-Nikon Imaging Centre for imaging assistance; the research coordination team and Department of Pathology at NUH for assistance with human samples; S. Sagiraju for assistance with animal experiments; A. Lin and A. Ng for empirical candidate validation; K. Saito for assistance with RNA-seq preparation; M. Taniguchi and K. Kita for assistance with FACS; D. H. Alpers for providing the GIF antibody; A. van Oudenaarden and A. Lyubimova for assistance with CEL-seq and RacelD; and F. de Sauvage for providing the *Lgr5-DTR-eGFP* mice. N.B. is supported by the Agency for Science, Technology and Research (A*Star), Singapore Gastric Cancer Consortium (SGCC) and Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number 17H01399. This research is supported by Singapore Ministry of Health's National Medical Research Council under its Open Fund-Young Individual Research Grant (NMRC/OFYIRG/0007/2016) and the National Research Foundation Singapore (Investigatorship Program award no. NRF-NRF12017-03).

Author contributions S.H.T. and Y.S. contributed to all aspects of the study: they designed, performed all empirical experiments, collected and analysed data, and wrote the manuscript. S.T. designed and performed experiments, and collected and analysed data for profiling and validation studies for candidate-marker identification. J.G. performed immunostaining, CEL-seq experiments and mouse husbandry. R.S. provided advice and technical help with human and mouse cancer, analysed data and wrote the manuscript. K.M. performed FACS and immunostaining for human AQP5 FACS experiments. P.P. performed immunostaining and mouse husbandry. L.T.T. performed mouse husbandry. E.W. generated the transgenic mouse lines. T.S. and S.W.T.H. analysed human cancer data in pathway analysis. S.L.I.J.D. analysed microarray, CEL-seq and RNA-seq data. S.M. performed FACS experiments. A.F. provided advice and technical help with human experiments and mouse cancer models. M.O., T.T., H.I.G., S.S., M.T., K.G.Y., J.S. and A.S. provided patient samples. H.I.G., S.S. and M.T. analysed and scored stained patient samples. P.T. designed and supervised cancer frequency analysis. N.B. supervised the project, analysed the data and wrote the manuscript. All authors discussed results and edited the manuscript.

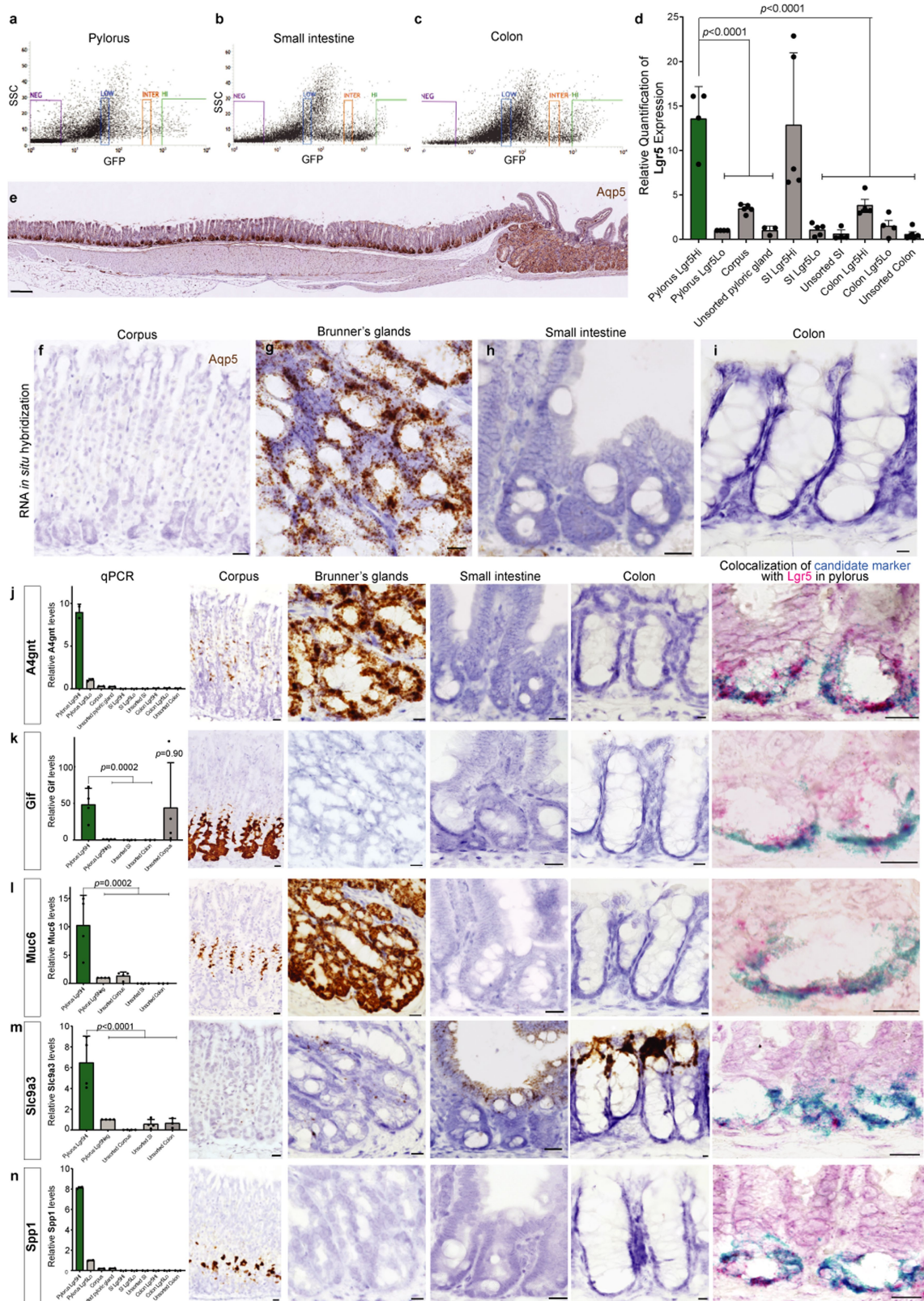
Competing interests N.B. and S.H.T. are co-inventors on the provisional patent application 10201911742W titled 'A method for functional classification and diagnosis of cancers'. This patent covers the analysis of human cancers using their signalling pathway statuses. All the other authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-1973-x>.

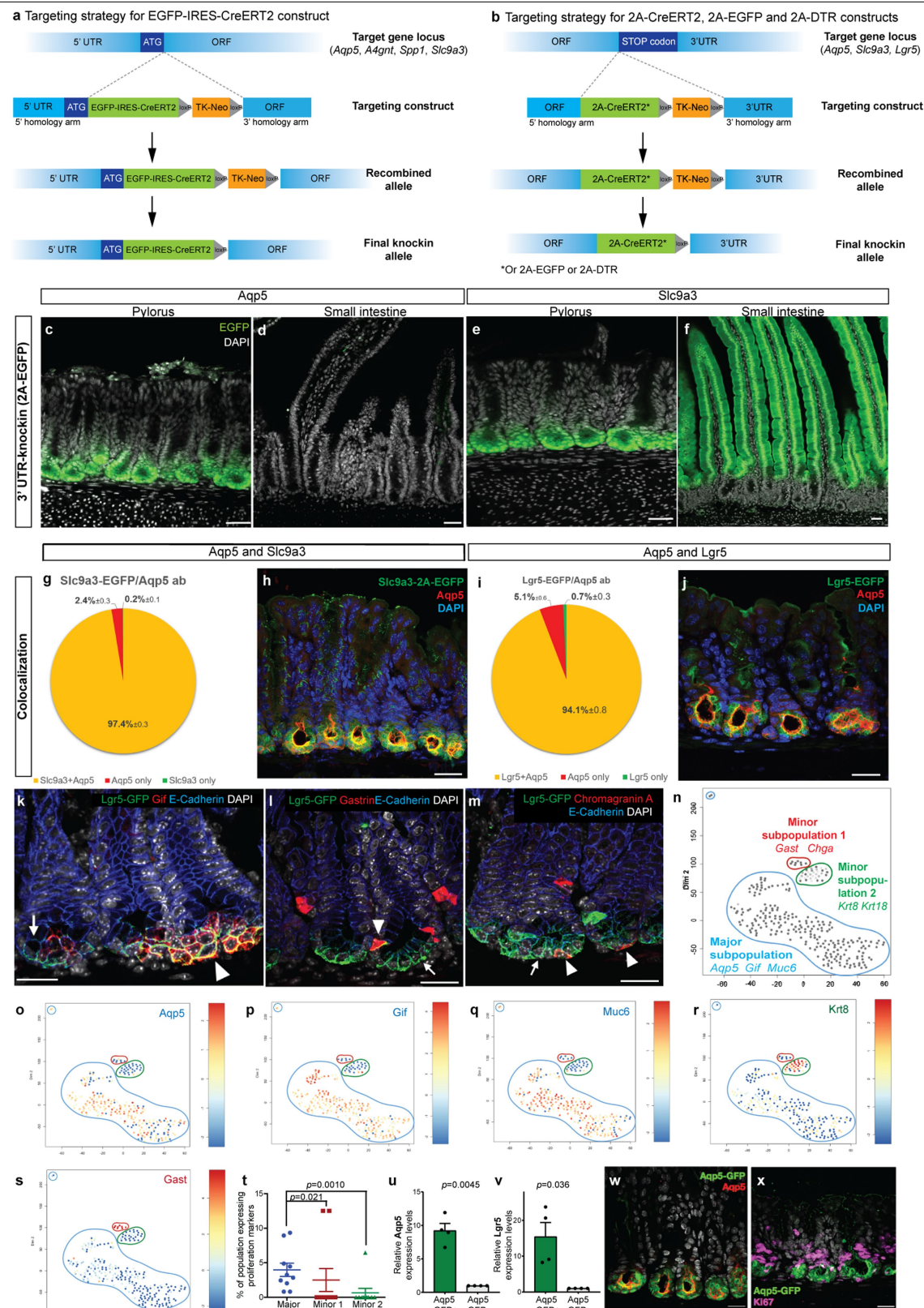
Correspondence and requests for materials should be addressed to N.B.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | Comparative profiling of LGR5 populations in gastrointestinal tissues identifies new pyloric-specific markers. **a–c**, FACS strategy sorting eGFP^{high} and eGFP^{low} cells from *Lgr5-eGFP-IRES-creERT2* pylorus (**a**), small intestine (**b**) and colon (**c**). **d**, *Lgr5* expression (by qPCR) in sorted populations and unsorted tissues of the gastrointestinal tract. Data are represented as mean \pm s.e.m., $n = 4$ biological replicates; one-way ANOVA. **e**, AQP5 protein expression in the mouse stomach through to the duodenum by immunostaining, $n = 3$ biological replicates. **f–i**, *Aqp5* mRNA expression in the

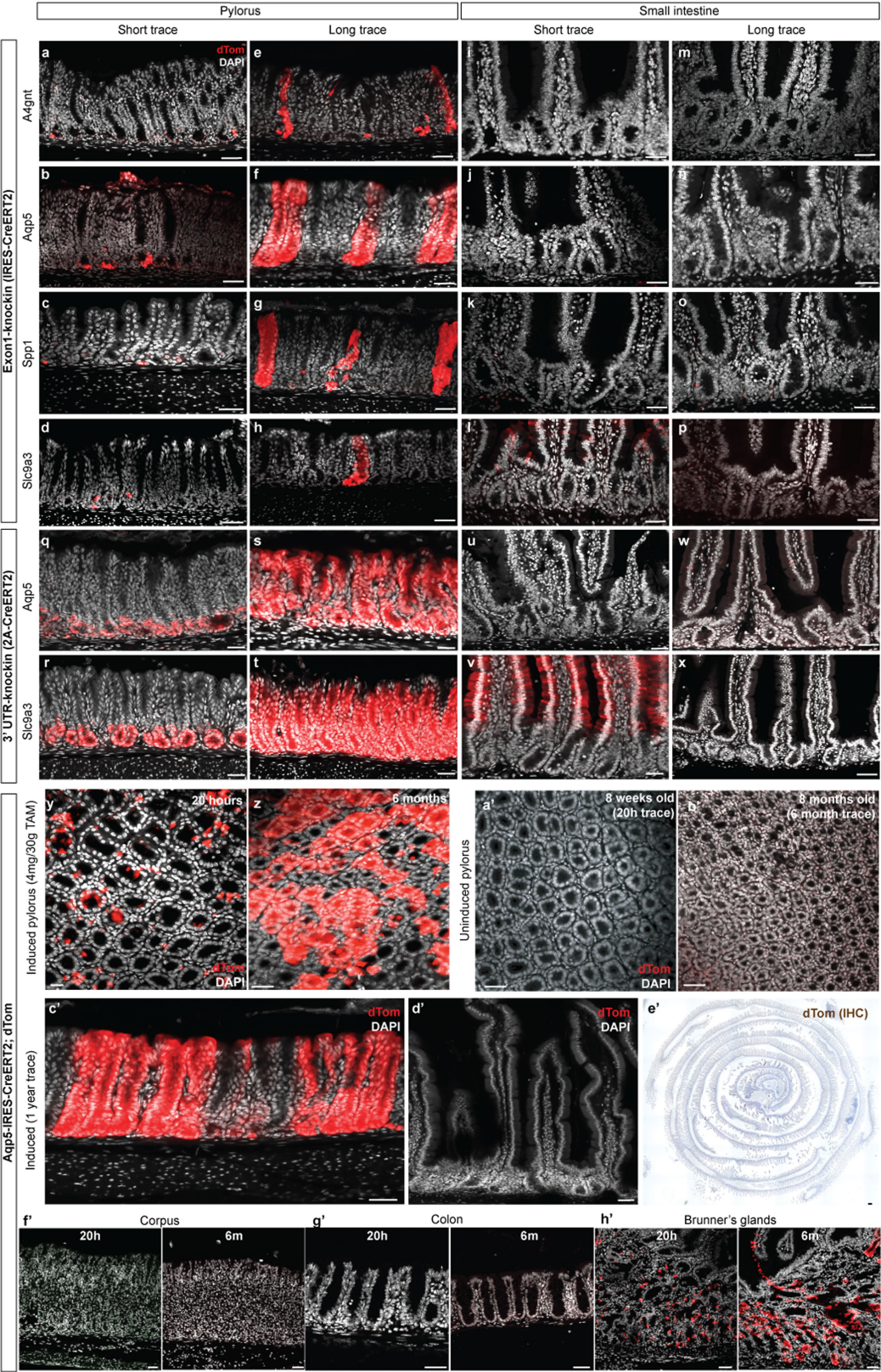
corpus (**f**), Brunner's glands (**g**), small intestine (**h**) and colon (**i**) by ISH. $n = 3$ biological replicates. **j–n**, *A4gnt* (**j**), *Gif* (**k**), *Muc6* (**l**), *Slc9a3* (**m**) and *Spp1* (**n**) expression in the corpus, Brunner's glands, small intestine and colon by qPCR, ISH and co-ISH with LGR5. For histology experiments, $n = 3$; for qPCR, $n = 4$ biological replicates for *Gif*, *Muc6* and *Slc9a3* qPCRs, for which data are represented as mean \pm s.e.m., $n = 2$ technical replicates from a pooled sample of 8 for *A4gnt* and *Spp1* qPCRs. Scale bars, 500 μ m (**e**), 20 μ m (**f–n**).



Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | AQP5 marks the major subpopulation of LGR5^{high} pyloric stem cells. **a**, Exon 1 knock-in gene strategy to generate EGFP-CreERT2 reporters of *Aqp5*, *A4gnt*, *Slc9a3* and *Spp1* expression. **b**, The 3' UTR knock-in gene strategy to generate 2A-eGFP, 2A-CreERT2 or 2A-DTR reporters of *Aqp5*, *Slc9a3* and *Lgr5* expression. **c-f**, eGFP signal in the pylorus and small intestine of *Aqp5*-2A-eGFP (**c, d**) and *Slc9a3*-2A-eGFP mice (**e, f**). **g, h**, Quantification of the overlap between eGFP⁺ cells and AQP5⁺ cells in *Slc9a3*-2A-eGFP pylori (**g**) ($n = 102$ glands from 3 mice) and a representative image of the immunostaining (**h**). Results are presented as mean \pm s.e.m. **i, j**, Quantification of overlap between LGR5-eGFP⁺ cells with AQP5⁺ cells (**i**) ($n = 117$ glands from 4 mice) and a representative image of the immunostaining (**j**). Results are presented as mean \pm s.e.m. **k-m**, Colocalization of LGR5-eGFP in the pylorus with GIF (**k**), GAST (**l**) and CHGA (**m**). $n = 3$ biological replicates. **n**, t -distributed stochastic

neighbour embedding (t -SNE) map of single LGR5-eGFP^{high} cells from the pylorus. $n = 285$ cells from 3 mice. **o-s**, t -SNE maps showing enrichment of candidate markers in major (**o-q**) and minor (**r, s**) subpopulations of LGR5^{high} pyloric cells. $n = 285$ cells from 3 mice. **t**, Frequency of 10 published proliferation markers (*Bcl2*, *Ccnd1*, *Ckap2*, *Foxm1*, *Ki67*, *Mcm2*, *Mybl2*, *Plk1*, *Rrm2* and *Top2a*) in major versus minor subpopulations, compared by two-tailed Mann-Whitney test. $n = 285$ cells from 3 mice, 248 cells in major, 8 cells in minor-1 and 29 cells in minor-2 populations. **u, v**, *Aqp5* (**u**) and *Lgr5* (**v**) expression in cells sorted from *Aqp5*-eGFP-IRES-creERT2 pylori. Mean \pm s.e.m., $n = 4$ biological replicates. **w, x**, Co-immunostaining for eGFP driven by *Aqp5*-eGFP-IRES-creERT2 and endogenous AQP5 (**w**) and Ki67 (**x**). $n = 3$ mice. Scale bars, 50 μ m (**g, h**), 25 μ m (**k, m, w, x**).

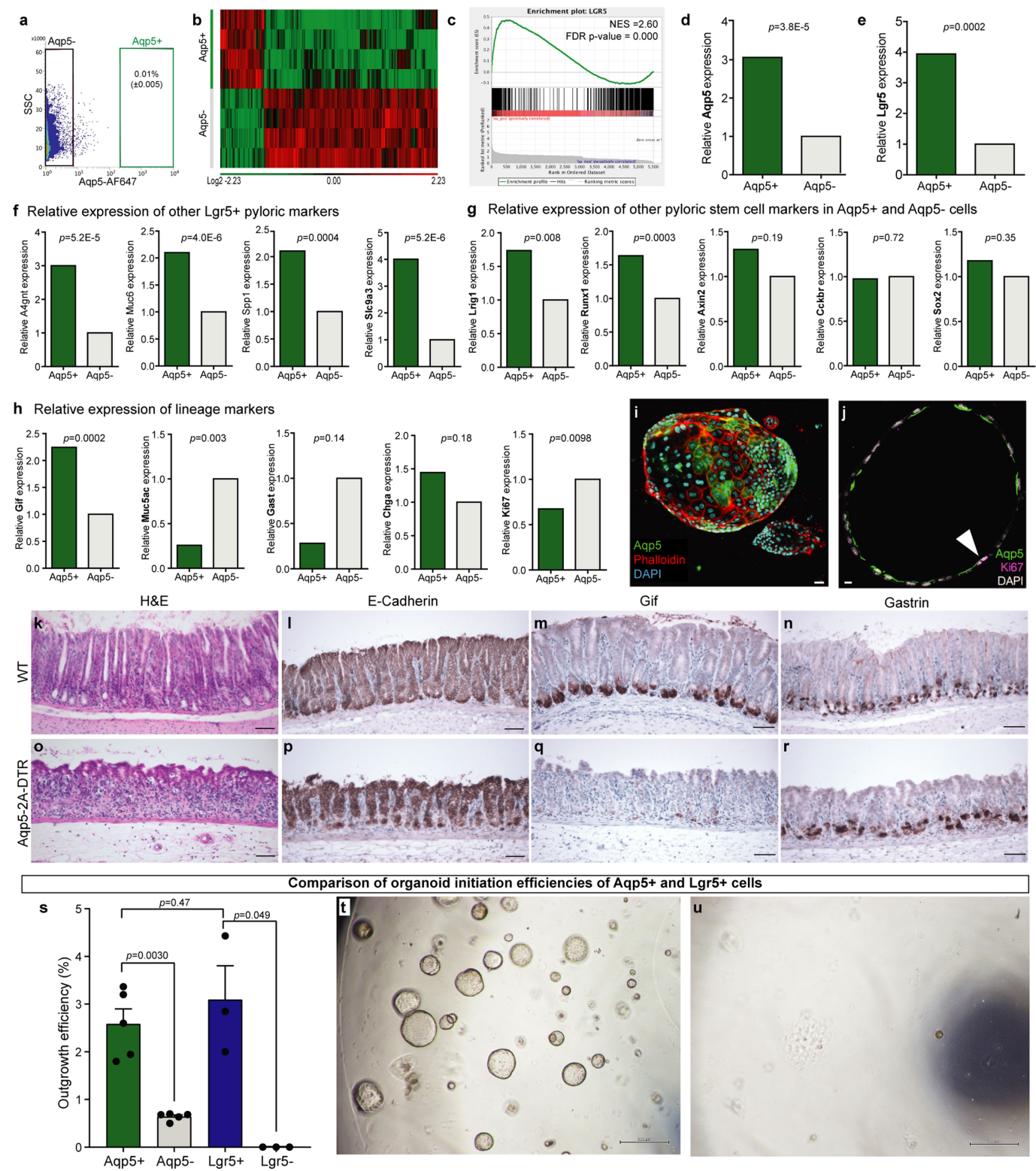


Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | AQP5 and other newly identified pyloric markers label pyloric stem cells, but not other gastrointestinal stem cells, in vivo.

a–p, Lineage tracing in *A4gnt*⁺, *Aqp5*⁺, *Spp1*⁺ and *Slc9a3-eGFP-IRES-creERT2* mice crossed with *tdTomato*^{LSL} reporter mice after a short trace (20–48 h) in the pylorus (**a–d**) and small intestine (**i–l**), and a long trace (>3 months) in the pylorus (**e–h**) and small intestine (**m–p**). *n* = 3 mice per genotype. **q–x**, Lineage tracing in pylorus (**q–t**) and small intestine (**u–x**) of the *Aqp5-2A-creERT2* and *Slc9a3-2A-creERT2* mice after a short trace (**q–r**, **u**, **v**) and a long trace (**s**, **t**, **w**, **x**). **y, z, a', b'**, Whole-mount imaging of pylorus from induced *Aqp5-eGFP-IRES-*

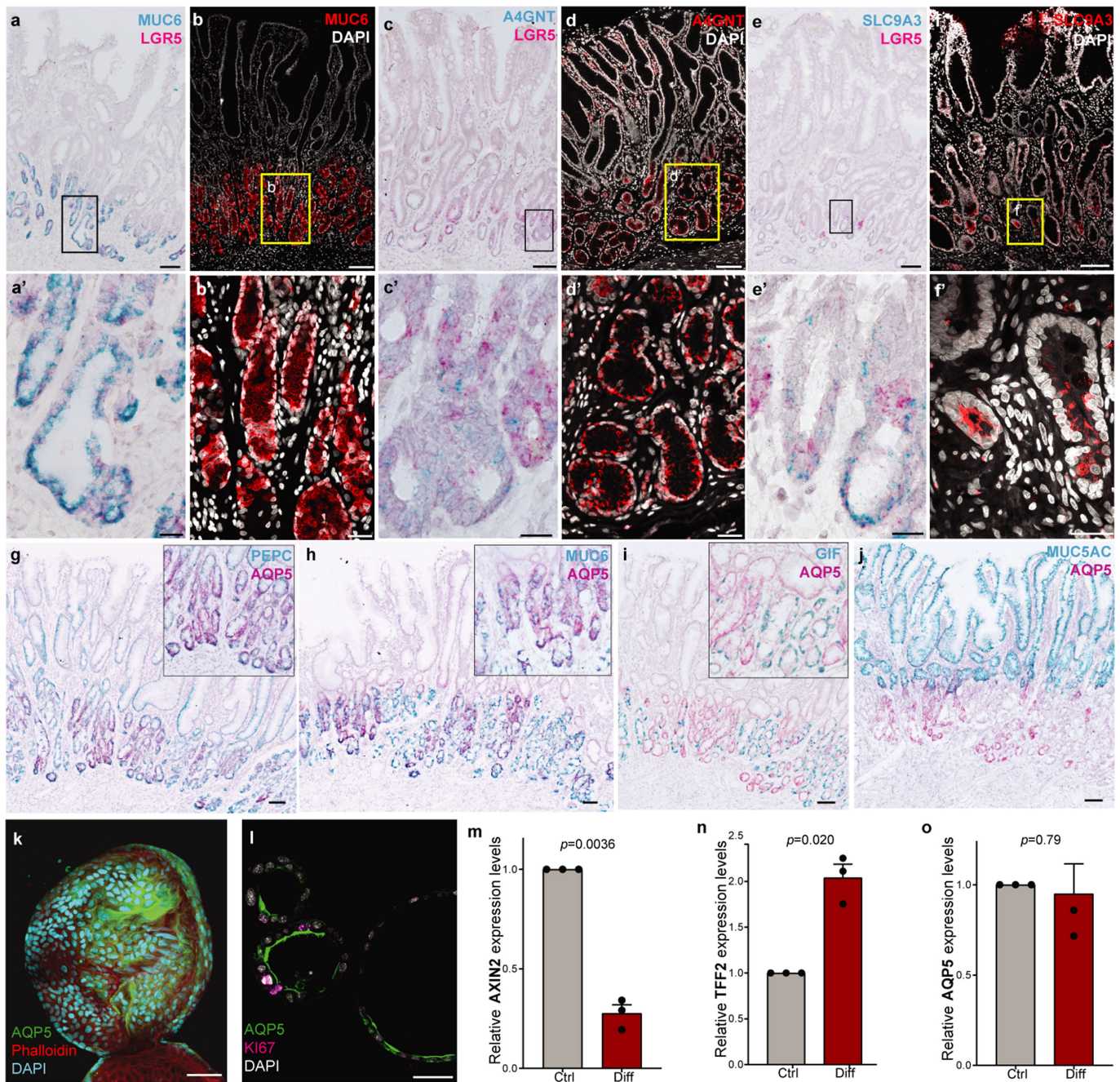
creERT2;tdTomato^{LSL} mice 20 h (**y**) and 6 months (**z**) after induction. Whole-mount imaging of pylorus from uninduced 8-week-old (**a'**) and 8-month-old (**b'**) *Aqp5-eGFP-IRES-creERT2;tdTomato*^{LSL} mice. *n* = 3 mice per condition. *tdTomato* (dTom) signal through the entire height of the pyloric epithelium is shown, and DAPI from the upper parts of pyloric glands is depicted for clarity. **c', d', e'**, *tdTomato* signal in clusters of glands in 1-year-traced pylorus (**c'**) and small intestine (**d', e'**). **f', g', h'**, *tdTomato* expression in gastric corpus (**f'**), colon (**g'**) and Brunner's glands (**h'**) 20 h and 6 months after induction. *n* = 3 biological replicates. Scale bars, 50 μ m.



Extended Data Fig. 4 | See next page for caption.

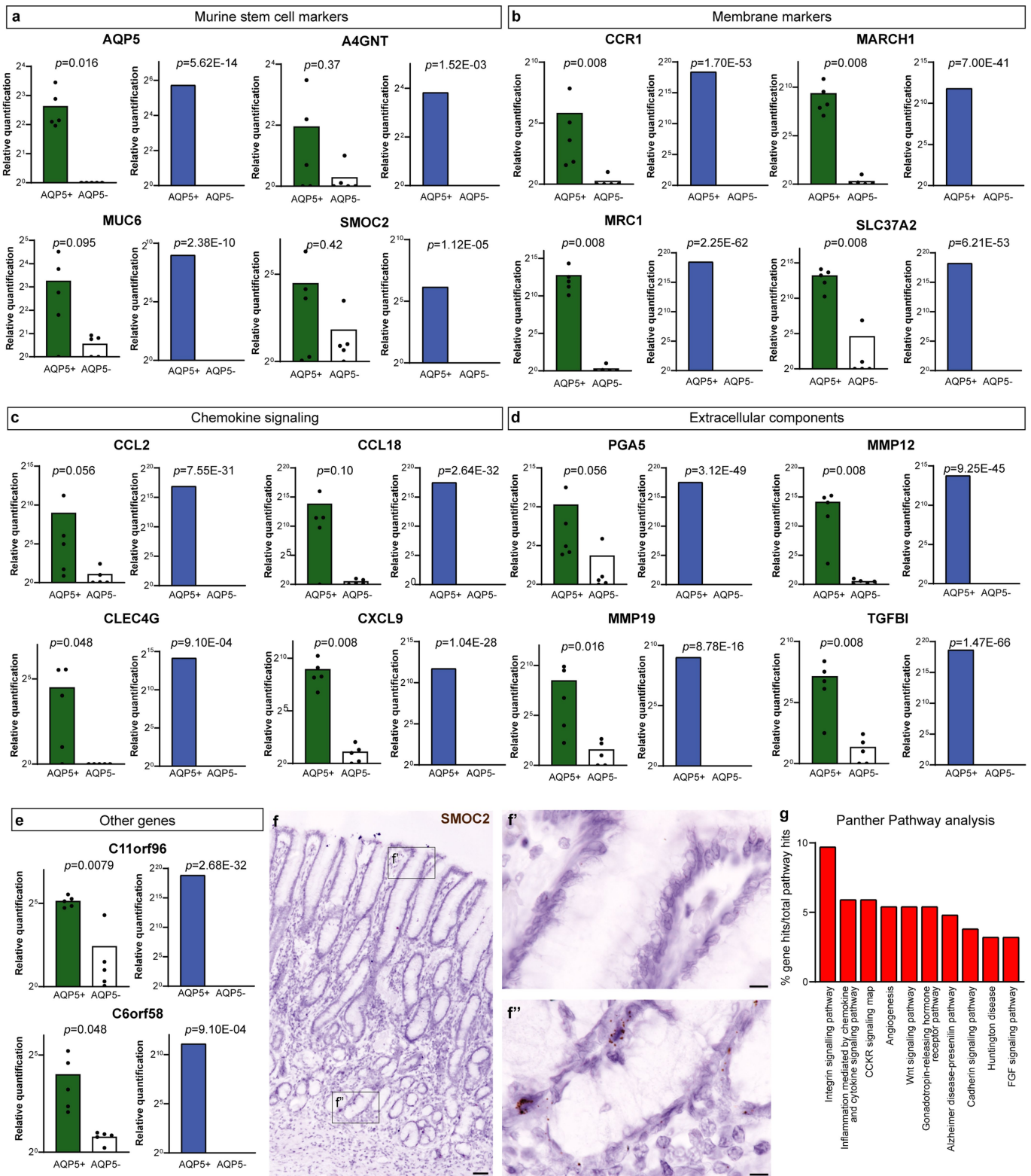
Extended Data Fig. 4 | Detailed characterization of AQP5-expressing pyloric cells by transcriptomic profiling, in vivo ablation and ex vivo organoid culture. **a**, AQP5⁺ gating strategy with cells from wild-type pylorus stained only with propidium iodide. **b**, Heat map of transcriptomes from AQP5⁺ and AQP5⁻ cells. *n* = 4 biological replicates. **c**, Gene set enrichment analysis comparing the degree of overlap between transcriptomes of AQP5⁺ cells and LGR5⁺ cells from the pylorus using Kolmogorov–Smirnov statistic. *n* = 4 biological replicates each. **d**, **e**, Relative *Aqp5* (**d**) and *Lgr5* (**e**) expression (from microarray) in AQP5⁺ and AQP5⁻ cells (*n* = 4 biological replicates), by one-way ANOVA in the Partek analysis software. **f–h**, Relative expression of various pyloric markers (**f**), other published pyloric stem cell markers (**g**) and lineage and proliferation markers (**h**) in AQP5⁺ population versus AQP5⁻ population in microarray. *n* = 4 biological

replicates. Data are represented as mean, as derived from Partek analysis software by one-way ANOVA. **i**, **j**, AQP5 staining in a whole-mount organoid (**i**) and AQP5 colocalization with KI67 (**j**) in an organoid section. Organoids were derived from single AQP5⁺ cells. *n* = 3 biological replicates. **k–r**, Pylori of diphtheria-toxin-treated wild-type (**k–n**) and *Aqp5-2A-DTR* (**o–r**) mice stained for H & E (**k**, **o**), E-cadherin (**l**, **p**), GIF (**m**, **q**) and GAST (**n**, **r**). *n* = 3 biological replicates. **s–u**, Outgrowth efficiency of AQP5⁺, AQP5⁻, LGR5⁺-eGFP⁺, LGR5⁻-eGFP⁺ cells (**s**) *n* = 5 biological replicates for AQP5⁺ and AQP5⁻ cells, *n* = 3 biological replicates for eGFP⁺ and eGFP⁻ cells from *Lgr5-2A-eGFP* pylori. Representative images of organoids derived from eGFP⁺ (**t**) and eGFP⁻ (**u**) cells from *Lgr5-2A-eGFP* pylori. Paired two-sided *t*-test. Scale bars, 25 μ m (**i**, **j**), 50 μ m (**k–r**), 500 μ m (**t**, **u**).



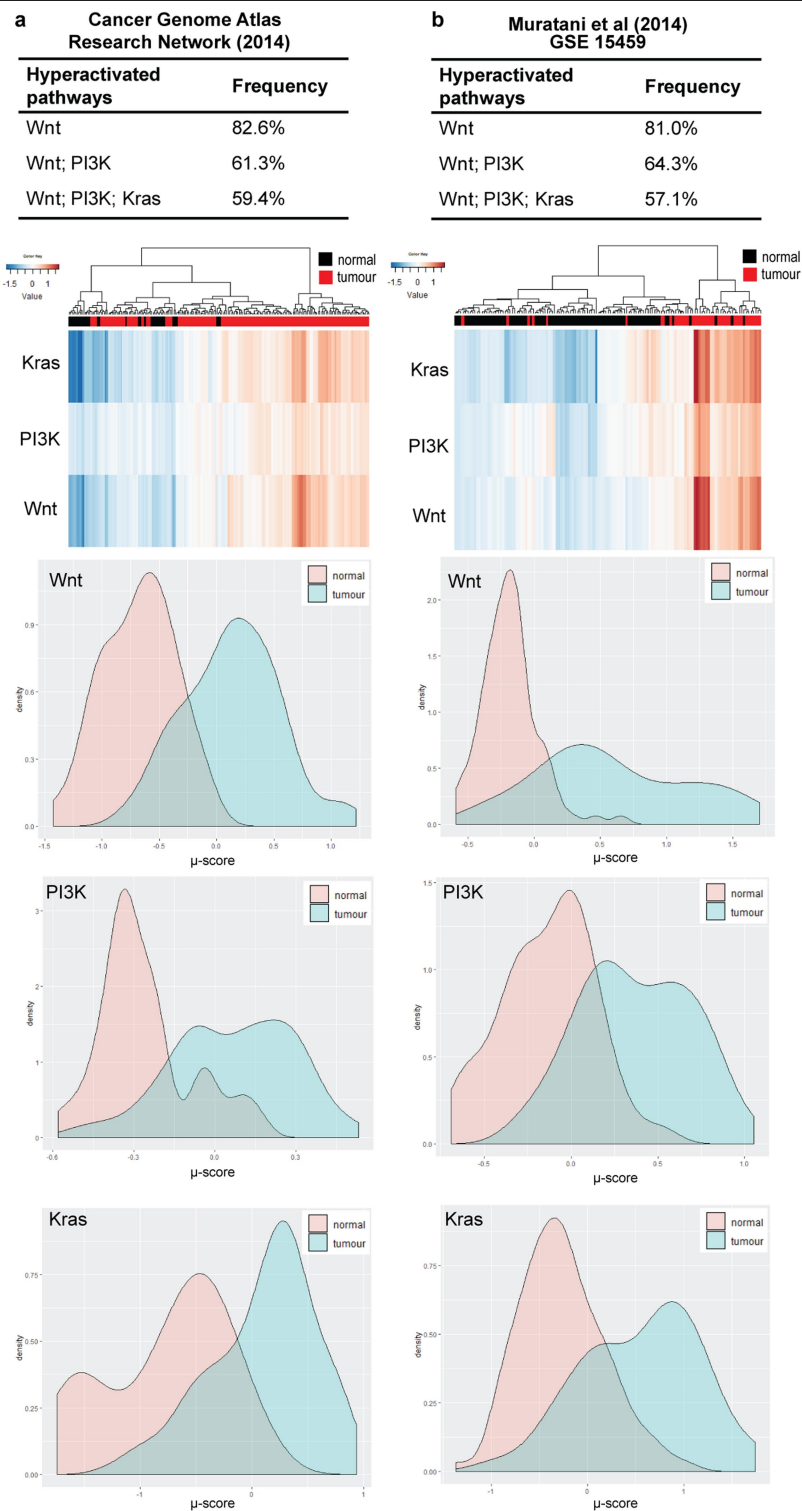
Extended Data Fig. 5 | AQP5 is expressed at human pyloric gland bases together with other pyloric markers, and facilitates the isolation of human pyloric stem cells. a–f, a', b', c', d', e', f', MUC6 (a, b, a', b'), A4GNT (c, d, c', d') and SLC9A3 (e, f, e', f') expression (co-ISH with LGR5 and immunostaining) in normal human pylorus. $n=3$ biological replicates. g–j, Co-ISH to colocalize AQP5 with PEPC (g), MUC6 (h), GIF (i) and MUC5AC (j) pyloric lineage markers.

$n=3$ biological replicates. k, l, AQP5 labelling in whole-mount human organoids (k) and AQP5 colocalization with Ki67 (l) in organoid sections. $n=3$ biological replicates. m–o, Relative AXIN2 (m), TFF2 (n) and AQP5 (o) expression in AQP5⁺ cell-derived organoids three days after WNT3A, Noggin and FGF10 withdrawal, by qPCR. $n=3$ biological replicates. Scale bars, 100 μ m (a–j, l), 25 μ m (a' b', c', d', e', f', k).



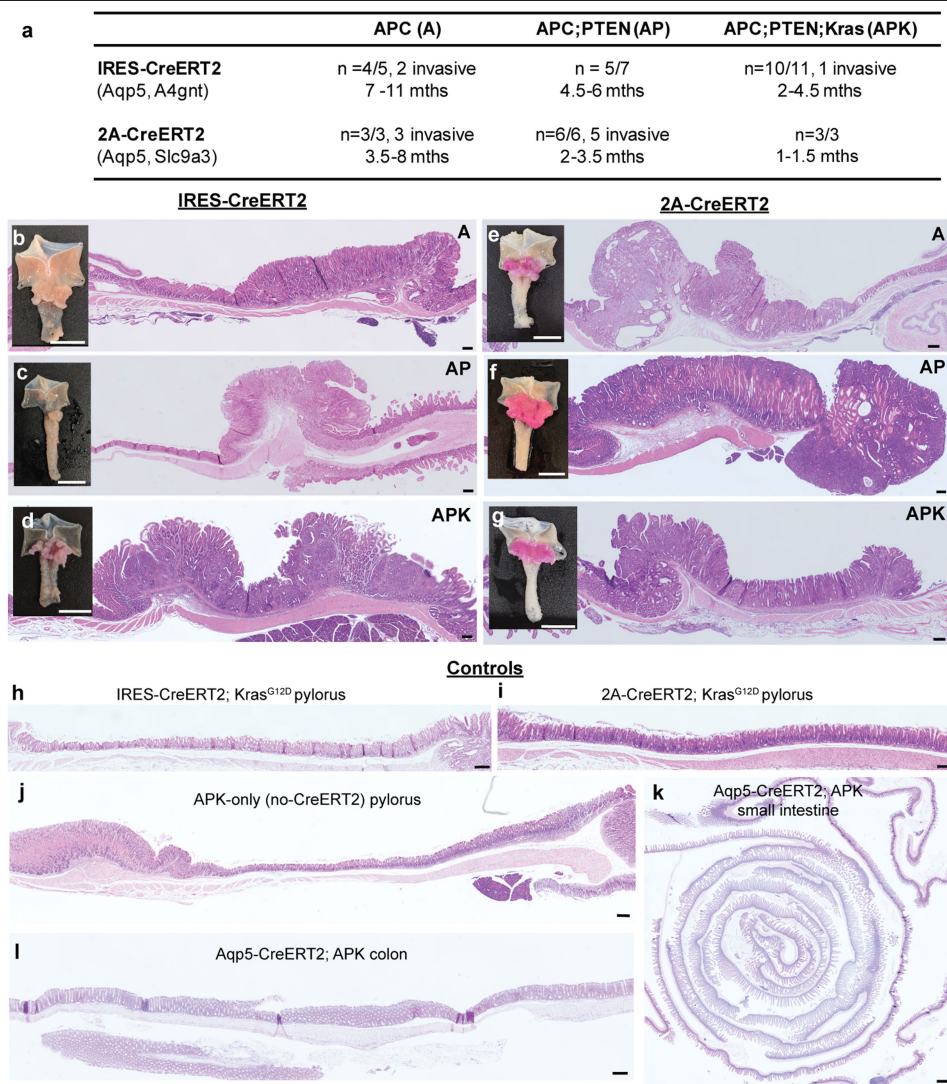
Extended Data Fig. 6 | Profiling and validation of the transcriptome of the human pyloric AQP5-expressing population. a–e, qPCR validation (green) ($n = 5$ biological replicates) and RNA-seq values (blue) ($n = 8$ biological replicates) of homologues of mouse stem cell markers (a), membrane components (b), chemokine signalling components (c), extracellular matrix components (d) and other genes (e). Two-sided Mann–Whitney test was used to determine statistical significance of qPCR result differences for all genes

except *AQP5*, which was determined by two-tailed paired *t*-test. qPCR and RNA-seq results are presented as mean. **f, f', f'',** IISH of *SMOC2* (brown) on normal human pylorus. **f'** is a magnified inset of surface mucosa, and **f''** is a magnified inset of gland base. $n = 4$ biological replicates. Scale bars, 100 μm (**f**), 10 μm (**f', f''**). **g,** Top 10 Panther pathways enriched with the most candidate genes.



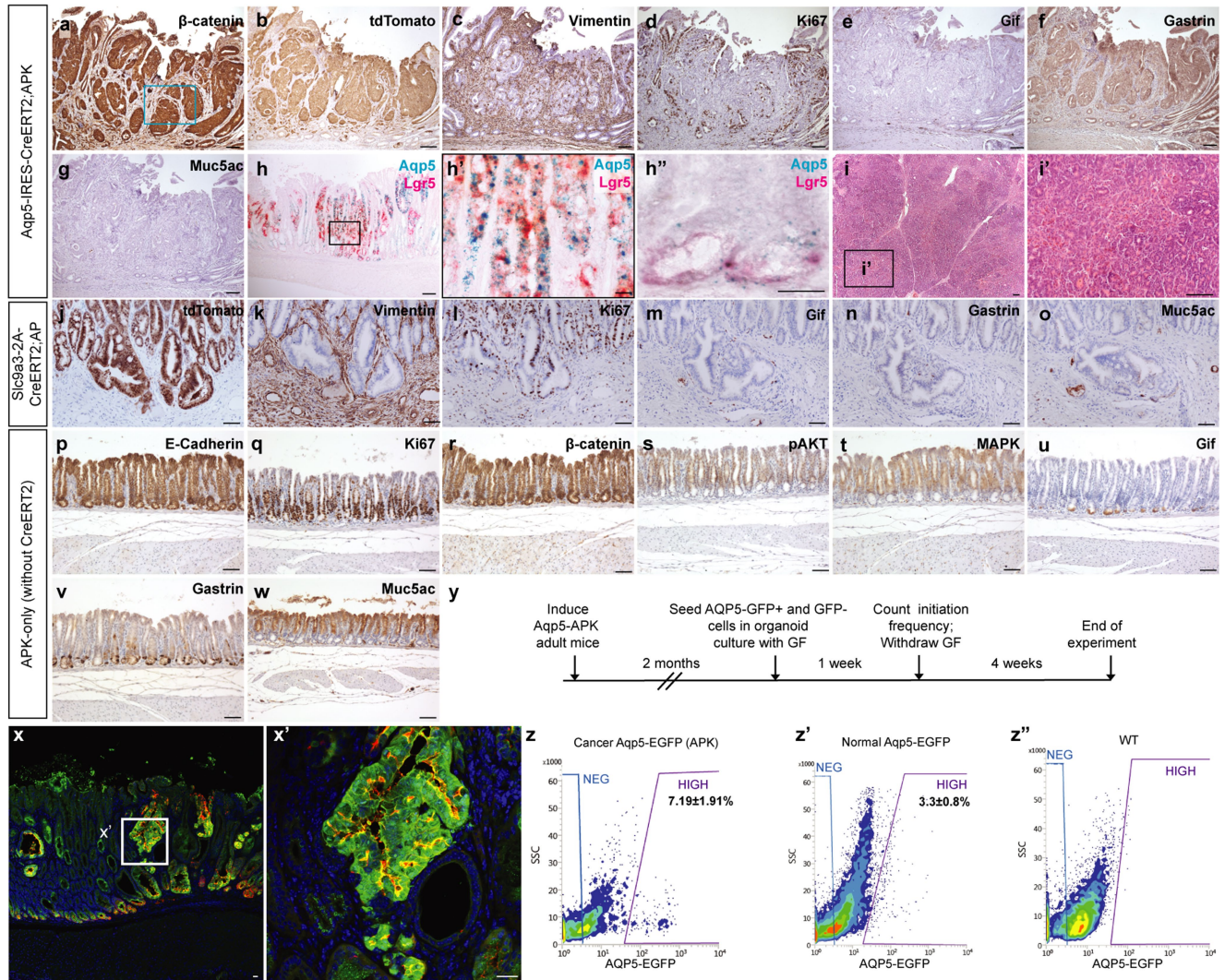
Extended Data Fig. 7 | WNT, PI3K and KRAS pathways are commonly co-dysregulated in human distal gastric cancers. a, b, Co-hyperactivation status of the WNT, PI3K and KRAS pathways in human distal gastric cancer samples from TCGA³ (**a**) (*n* = 155) and GSE15459¹⁶ (**b**) (*n* = 42) datasets. Heat maps show

distribution of pathway hyperactivation status across samples. Graphs depict distribution of μ scores (degree of signalling activity) of normal and tumour samples for each of the pathways we examined.



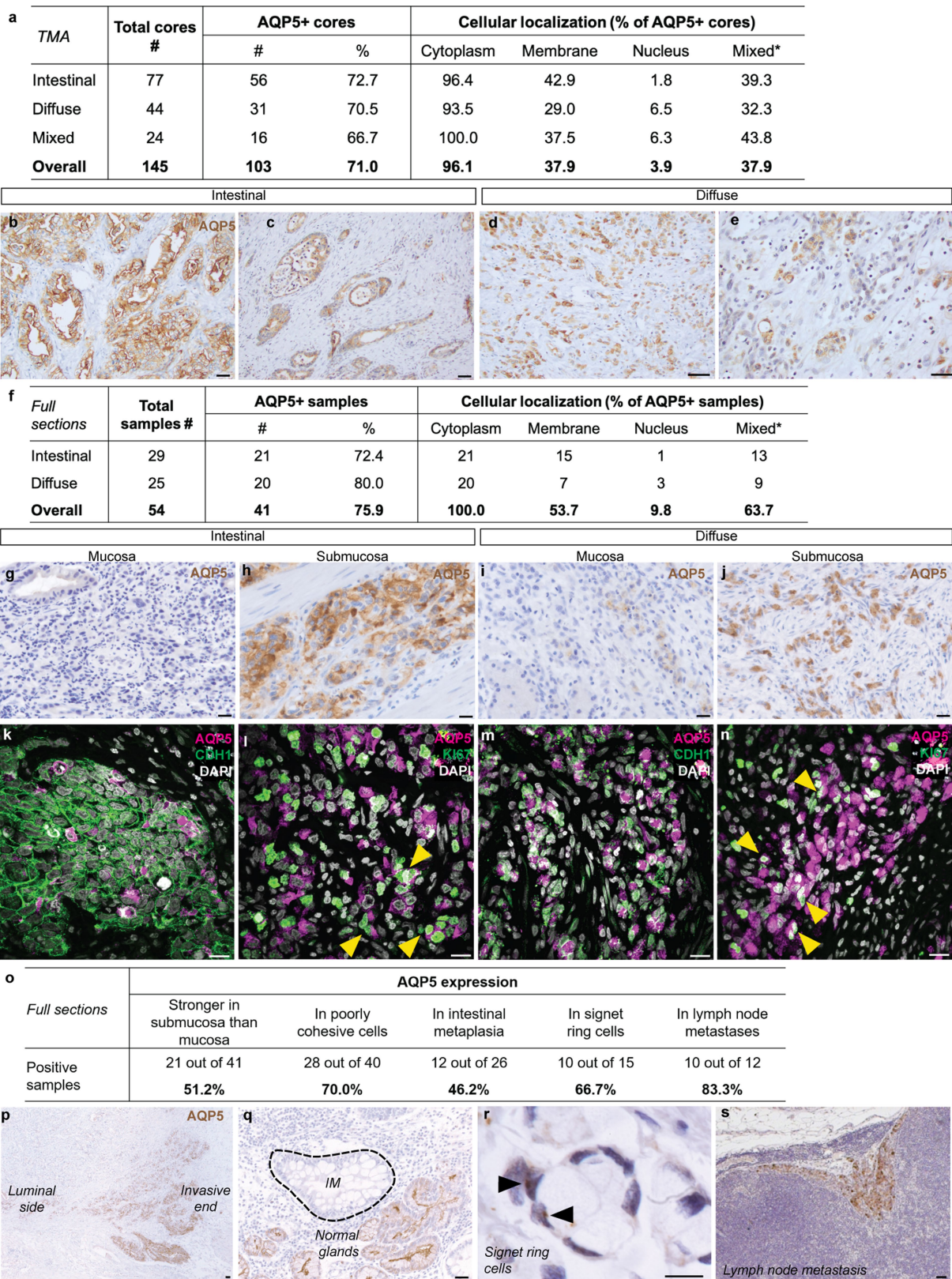
Extended Data Fig. 8 | Targeted conditional mutation of pyloric stem cells using our mouse models selectively drives tumour formation in the distal stomach. **a**, Sample sizes, tumour and invasion incidences observed in various permutations of *creERT2* drivers and oncogenic alleles. **b–g**, Whole-mount and H & E images of entire pyloric regions for each *creERT2*-oncogenic-allele

combination. **h–l**, H & E images of pylori from multiple-pyloric-marker-*creERT2*;Kras^{LSL-G12D} models (**h, i**), APK-only model (without *creERT2* driver) (**j**), and small intestine (**k**) and colon (**l**) from *Aqp5-IRES-creERT2* APK mouse model of gastric cancer. Scale bars, 1 cm (whole-mount insets in **b–g**), 200 μ m (H & E images in **b–l**).



Extended Data Fig. 9 | Phenotypic characterization of *Aqp5-IRES-creERT2* APK and *Slc9a3-2A-creERT2* AP distal stomach tumours. **a–g**, Immunostaining of various markers in *Aqp5-IRES-creERT2* APK pyloric tumour. **h, h', h''**, Co-ISH of AQP5 and LGR5 in tumour region; region in black box in **h** is magnified in **h'**. Dual ISH of AQP5 and LGR5 in an adjacent normal pyloric region from the same mouse (**h''**). **i, i'**, Representative H & E stain of a salivary gland tumour from *Aqp5-IRES-creERT2* APK mouse. **j–o**, Immunostaining of various markers in *Slc9a3-2A-creERT2* AP pyloric tumour. **p–w**, Immunostaining of various markers in the APK-only (no *creERT2*)

control pyloric stomach. **x, x'**, Colocalization of eGFP and endogenous AQP5 protein in pyloric tumour from an *Aqp5-IRES-creERT2* APK mouse. **y, z**, Control FACS gating for eGFP⁺ cells using normal *Aqp5-eGFP-IRES-creERT2* (y) and wild-type (z) pylori. *n* = 3 biological replicates. **y, z, z', z''**, Organoid assay for stemness of AQP5-eGFP⁺ tumour cells. *n* = 3 biological replicates. **y**, Experimental timeline. FACS gating strategy to isolate eGFP⁺ tumour cells from *Aqp5-IRES-creERT2* APK pyloric tumour (z), eGFP⁺ cells from normal *Aqp5-eGFP-IRES-creERT2* pylorus (z'), and control GFP gating with wild-type pylorus (z''). Scale bars, 100 μ m (a–g, i–x, h', i', x'), 20 μ m (h'').



Extended Data Fig. 10 | AQP5 expression is commonly dysregulated in human distal gastric cancers. **a**, Summary of AQP5 expression in a tumour microarray panel of 145 cores of human distal gastric cancer. AQP5 expression is scored as positive if observed in >5% of malignant cells. **b–e**, Examples of AQP5+ cores with intestinal (**b, c**) and diffuse (**d, e**) subtypes, often with cytoplasmic and/or membranous staining. **f**, Summary of AQP5 expression from 54 full sections of distal human gastric cancer. *Mixed refers to AQP5 localization in cytoplasm and nucleus, or cytoplasm and membrane. **g–n**, AQP5

expression in intestinal (**g, h, k, l**) and diffuse (**i, j, m, n**) subtypes. Yellow arrowheads indicate cells co-expressing AQP5 and Ki67. **o**, Summary of other observations of AQP5 expression in full gastric tumour sections. **p–s**, AQP5 expression in the invasive edge of the tumour (**p**), intestinal metaplasia (IM) (**q**) (dotted lines denote intestinal metaplasia region that is negative for AQP5), Signet ring cells (**r**) (black arrows denote cells with weak AQP5 expression) and tumour cells in lymph node metastasis (**s**). Scale bars, 20 μ m (**g–n, q, r**), 50 μ m (**b–e, p, s**).

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- ☒ ☐ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☒ ☐ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☒ ☐ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated
- ☐ ☒ Clearly defined error bars
State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on [statistics for biologists](#) may be useful.

Software and code

Policy information about [availability of computer code](#)

Data collection

BD FACS Software sorter software (v1.1), QuantStudio Real-Time PCR Software (v1.3), Nikon NIS-Elements AR (v5.11), Zeiss ZEN Blue (v2), Olympus FV31-SW (v2.3), Olympus DP2-BSW (v2.2)

Data analysis

Software: Partek Genomics Suite, GraphPad Prism (v5.03 and v8.1) and ImageJ (v1.52a), Bitplane (v8.0), Nikon NIS-Elements AR (v5.11) QuantStudio Real-Time PCR Software (v1.3)
Open source code: RaceID (<https://github.com/dgrun/RaceID>), edgeR version 3.12.1 (with limma_3.26.9), PANTHER (v14.1), GSEA (v6.1)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The datasets generated during and/or analysed during the current study are available in the GEO repository under accession codes GSE121803 and GSE133036. The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculation was performed. All qualitative experiments (e.g. immunostaining) were done with at least three biological replicates with all showing the same outcome, indicating applicability to the broader sample. We included as many independent replicates as possible (3 or more) in quantitative experiments, and used statistical tests to determine if any observable difference was statistically significant. Sample size is deemed sufficient with a statistically significant difference.
Data exclusions	In principle, data were only excluded for failed experiments resulting from technical error.
Replication	We performed experiments on at least 3 biological replicates to ensure reproducibility of results.
Randomization	No randomization of mice. Mice analyzed were litter mates and sex-matched whenever possible.
Blinding	Investigators were not blinded to mouse genotypes and patient conditions during experiments. Data reported for mouse experiments are not subjective but based on experimental observations. Investigators needed knowledge of patients' conditions (healthy or cancerous) to ascertain applicability of human tissues collected for downstream experiments.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Unique biological materials
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Unique biological materials

Policy information about [availability of materials](#)

Obtaining unique materials Unique mouse models are available from corresponding author at requests.

Antibodies

Antibodies used Detailed antibody information as requested has been compiled in Supplementary Table 6. The following is a list of antibodies

used:

chicken anti-EGFP (1:2,000; Abcam, ab290), rabbit anti-EGFP (1:200; Cell Signalling, 2956S), rabbit anti-Ki67 (1:200; ThermoFisher, MA5-14520), rabbit anti-GIF (1:10,000; provided by D. H. Alpers, Washington University School of Medicine, USA), rabbit anti-RFP (1:200; Rockland, 600-401-379), rabbit anti-aquaporin5 (1:200; Santa Cruz, SC-28628), rabbit anti-Slc9a3 (1:200; Santa Cruz, SC-16103-R), rabbit anti-mucin6 (1:200; LsBio, LS-C312108), rabbit anti-A4gnt (1:500; Novus Biologicals, NBP1-89129), rabbit anti-Gastrin (1:200; Novocastra, NCL-GASp), mouse anti-MUC5AC (1:200; Novocastra, NCL-HGM-45-M1), rabbit anti-vimentin (1:500; Abcam, ab92547), mouse anti-E-cadherin (1:200; BD Transduction Laboratories, 610181), mouse anti- β -catenin (1:200; BD Transduction Laboratories, 610154), mouse anti-RFP (1:200; Abcam, 129244), mouse anti-ChgA (1:200; Abcam, 15160), rabbit anti-Phospho-MAPK (1:200; Cell Signalling, 4370S), mouse anti-H-K-ATPase (1:1,000; MBL International Corporation, D032-3), rabbit anti-Phospho-Akt (1:200; Cell Signalling, 3787L)

Validation

The antibodies were validated by the relevant companies and show expected staining patterns and cellular localization in our experiments. The requested details have been summarised in Supplementary Table 6.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

Adult B6 mice (7 weeks and older) of both genders were used in the study. All animal experiments were approved by the Institutional Animal Care and Use Committee of Singapore.

Wild animals

The study did not involve wild animals.

Field-collected samples

The study did not involve samples collected from the field.

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics

Healthy gastric mucosa or gastric adenocarcinomas from patients' pyloric antrum were collected for the study. State of health was determined by a clinical pathologist. Healthy mucosa was collected in Advanced DMEM/F12 media for processing in lab, while gastric adenocarcinomas were collected in 4% paraformaldehyde for FFPE processing.

Recruitment

Participants were recruited by the National University Hospital research coordination team, and provided by National University Singapore Department of Medicine and Pathology (granted under protocol-11-167E). Informed consent was obtained from all patients.

Flow Cytometry

Plots

Confirm that:

- ☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☒ All plots are contour plots with outliers or pseudocolor plots.
- ☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Murine pylorus was incubated in chelation buffer (5.6 mM sodium phosphate, 8 mM potassium phosphate, 96.2 mM sodium chloride, 1.6 mM potassium chloride, 43.4 mM sucrose, 54.9 mM D-sorbitol, 1 mM dithiothreitol) with 5 mM EDTA at 4°C for 2 hours. Glands were isolated by repeated pipetting of finely chopped pylorus tissue in cold chelation buffer. Chelation buffer containing isolated glands was filtered through 100 μ m filter mesh, and centrifuged at 720g at 4°C for 3 min. The pellet was resuspended in TrypLE (Life Technologies) with DNaseI (0.8U/ μ L)(Sigma) and incubated at 37°C for 10 min with intermittent trituration for digestion into single cells. Digestion was quenched by dilution with cold HBSS buffer. The suspension was centrifuged at 720g at 4°C for 3 min. For anti-Aqp5 antibody stain, the pellet was resuspended in HBSS with 2% fetal bovine serum (FBS, Hyclone) with anti-Aqp5-AF647 (Abcam) at 1:500 dilution and incubated on ice at 30min in the dark. The pellet was subsequently washed twice with cold HBSS and spun at 800g for 3min at 4°C. The pellet was resuspended in HBSS with 2% fetal bovine serum (FBS, Hyclone). Before sorting, 1 μ g/ml propidium iodide (Life Technologies) was added to the cell suspensions, filtered through a 40 μ m strainer, and sorted on BD Influx Cell Sorter (BD Biosciences).

Instrument

BD Influx

Software

BD FACSDiva for collection and population analysis

Cell population abundance

Sorts with GFP-reporter mice required one pylorus per sample. Sorts using antibodies required four pylori. Depending on the marker, the yield ranged from 1000 -3000 positive cells per experiment. Negative population is always in abundance. Cell populations collected were subsequently confirmed by qPCR.

Gating strategy

Positive/Negative gating strategy was defined with wildtype or unstained cells, and subsequently confirmed by qPCR.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

GDF15 mediates the effects of metformin on body weight and energy balance

<https://doi.org/10.1038/s41586-019-1911-y>

Received: 1 July 2019

Accepted: 16 December 2019

Published online: 25 December 2019

There are amendments to this paper

Anthony P. Coll^{1,10*}, Michael Chen², Pranali Taskar², Debra Rimmington¹, Satish Patel¹, John A. Tadross¹, Irene Cimino¹, Ming Yang¹, Paul Welsh³, Samuel Virtue¹, Deborah A. Goldspink¹, Emily L. Miedzybrodzka¹, Adam R. Konopka⁴, Raul Ruiz Esponda⁴, Jeffrey T.-J. Huang⁵, Y. C. Loraine Tung¹, Sergio Rodriguez-Cuenca¹, Rute A. Tomaz⁶, Heather P. Harding⁷, Audrey Melvin¹, Giles S. H. Yeo¹, David Preiss⁸, Antonio Vidal-Puig¹, Ludovic Vallier⁶, K. Sreekumaran Nair⁴, Nicholas J. Wareham⁹, David Ron⁷, Fiona M. Gribble¹, Frank Reimann¹, Naveed Sattar^{3,10}, David B. Savage^{1,10}, Bernard B. Allan^{2,10} & Stephen O'Rahilly^{1,10*}

Metformin, the world's most prescribed anti-diabetic drug, is also effective in preventing type 2 diabetes in people at high risk^{1,2}. More than 60% of this effect is attributable to the ability of metformin to lower body weight in a sustained manner³. The molecular mechanisms by which metformin lowers body weight are unknown. Here we show—in two independent randomized controlled clinical trials—that metformin increases circulating levels of the peptide hormone growth/differentiation factor 15 (GDF15), which has been shown to reduce food intake and lower body weight through a brain-stem-restricted receptor. In wild-type mice, oral metformin increased circulating GDF15, with *GDF15* expression increasing predominantly in the distal intestine and the kidney. Metformin prevented weight gain in response to a high-fat diet in wild-type mice but not in mice lacking GDF15 or its receptor GDNF family receptor α -like (GFRAL). In obese mice on a high-fat diet, the effects of metformin to reduce body weight were reversed by a GFRAL-antagonist antibody. Metformin had effects on both energy intake and energy expenditure that were dependent on GDF15, but retained its ability to lower circulating glucose levels in the absence of GDF15 activity. In summary, metformin elevates circulating levels of GDF15, which is necessary to obtain its beneficial effects on energy balance and body weight, major contributors to its action as a chemopreventive agent.

Metformin has been used as a treatment for type 2 diabetes since the 1950s. Recent studies have shown that it can also prevent or delay the onset of type 2 diabetes in people at high risk^{1,2}. At-risk individuals treated with metformin exhibit a reduction in body weight, glucose and insulin levels and enhanced insulin sensitivity³. Although many mechanisms for the insulin-sensitizing actions of metformin have been proposed⁴, they do not explain the weight loss. The robustness and persistence of metformin-induced weight loss in participants in the Diabetes Prevention Program has drawn attention to its importance to the chemopreventive effects of the drug⁵. A recent observational epidemiological study⁶ noted a strong association of metformin use with circulating levels of GDF15, a peptide hormone produced by cells responding to stressors⁷. GDF15 acts through a receptor complex that is expressed solely in the hindbrain, through which it suppresses food intake^{8–11}. We hypothesized that the effects of metformin in

lowering body weight may involve the elevation of circulating levels of GDF15.

Human studies

We first measured circulating GDF15 in a short-term human study¹² and found that, after two weeks of metformin treatment, there was an increase of about 2.5-fold in mean circulating GDF15 (Fig. 1a). To determine whether this increase was sustained, we measured circulating GDF15 levels at 6, 12 and 18 months in all available participants in carotid atherosclerosis: metformin for insulin resistance (CAMERA)¹³, an 18-month randomized placebo-control trial of metformin in people without diabetes but with a history of cardiovascular disease. In this study, metformin-treated participants lost about 3.5% of body weight with no significant change in weight in the placebo arm¹³. Metformin

¹MRC Metabolic Diseases Unit, Wellcome Trust-Medical Research Council Institute of Metabolic Science, University of Cambridge, Cambridge, UK. ²NGM Biopharmaceuticals, South San Francisco, CA, USA. ³Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK. ⁴Division of Endocrinology, Mayo Clinic, Rochester, MN, USA. ⁵Division of Systems Medicine, School of Medicine, University of Dundee, Dundee, UK. ⁶Wellcome-Medical Research Council Cambridge Stem Cell Institute, Department of Surgery, University of Cambridge, Cambridge, UK. ⁷Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK. ⁸MRC Population Health Research Unit, Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK. ⁹MRC Epidemiology Unit, Wellcome Trust-Medical Research Council Institute of Metabolic Science, University of Cambridge, Cambridge, UK. ¹⁰These authors contributed equally: Anthony P. Coll, Naveed Sattar, David B. Savage, Bernard B. Allan, Stephen O'Rahilly. *e-mail: apc36@cam.ac.uk; so104@medschl.cam.ac.uk

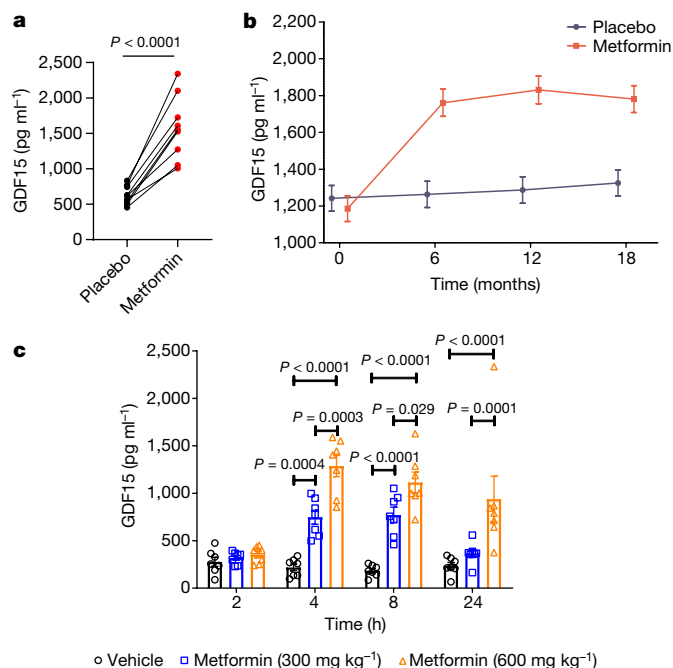


Fig. 1 | Effect of Metformin on circulating GDF15 levels in humans and mice. **a**, Paired serum GDF15 concentration in nine human subjects after two weeks of either placebo or metformin treatment, P value (95% CI) by two-tailed t -test. **b**, Plasma GDF15 concentration in overweight or obese non-diabetic participants with known cardiovascular disease, randomized to metformin or placebo in CAMERA, using a mixed linear model. Data are mean \pm s.e.m. Subject numbers: placebo and metformin, respectively, at time points: baseline, $n = 85$ and 86; 6 months, $n = 81$ and 71; 12 months, $n = 77$ and 68; 18 months, $n = 83$ and 74. Comparing metformin vs placebo groups, two-sided $P = 0.311$ at baseline, and $P < 0.0001$ at 6, 12 and 18 months individually. **c**, Serum GDF15 levels (mean \pm s.e.m.) in obese mice measured 2, 4, 8 or 24 h after a single oral dose of 300 mg kg $^{-1}$ or 600 mg kg $^{-1}$ metformin, $n = 7$ per group, P values by 2-way ANOVA with Tukey's correction for multiple comparisons.

treatment was associated with significantly ($P < 0.0001$) increased levels of circulating GDF15 at all three time points (Fig. 1b, Extended Data Fig. 1b–e). Furthermore, the change in serum GDF15 from baseline in metformin recipients was significantly correlated ($R = -0.26$, $P = 0.024$) with weight loss (Extended Data Fig. 1a).

The correlation of GDF15 increment with changes in body weight, while statistically significant, was modest in size. Although we believe that it does contribute to weight loss in some individuals taking metformin, we acknowledge that it is not necessary, and there are individuals with increases in GDF15 that do not exhibit weight loss. However, in the context of a long-term human study with imperfect drug compliance and intermittent sampling of GDF15 levels, it is noteworthy that such an association was seen at all. Further, there was no association of weight change with change in GDF15 in the placebo group ($R = -0.04$, $P = 0.740$, $n = 81$).

Mouse studies

Following these findings in humans, we performed a series of animal experiments to determine the potential causal link between the changes in GDF15 and weight changes induced by metformin. We administered metformin by oral gavage to mice fed a high-fat diet and measured serum GDF15. A single dose of 300 mg kg $^{-1}$ of metformin increased GDF15 levels for at least 8 h (Fig. 1c). A higher dose of metformin, 600 mg kg $^{-1}$, resulted in a sixfold increase in serum GDF15 levels at 4 h and 8 h after the dose, which were sustained above those of vehicle-treated mice for 24 h. The effects of metformin in chow-fed mice were

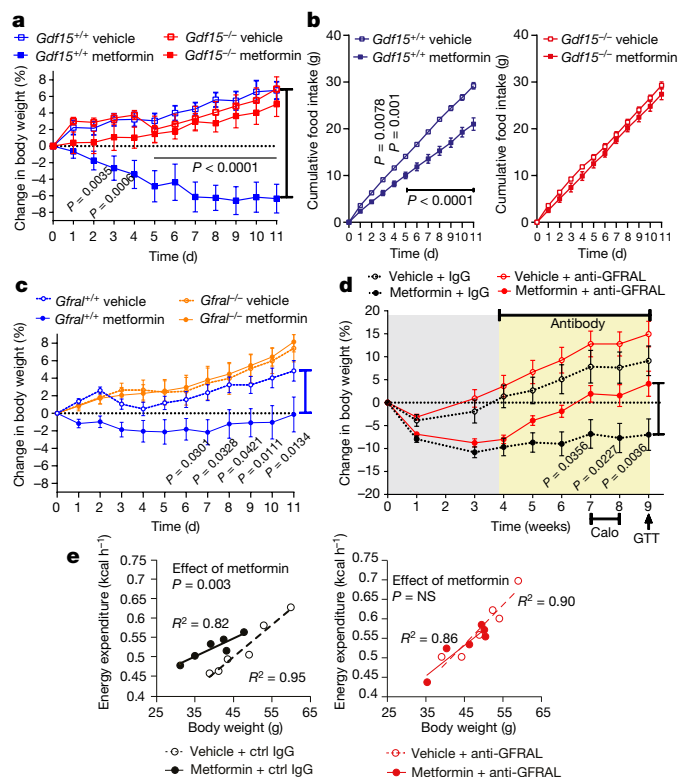


Fig. 2 | GDF15–GFRAL signalling is required for the weight-loss effects of metformin on a high-fat diet. **a**, Percentage change in body weight of $Gdf15^{+/+}$ and $Gdf15^{-/-}$ mice on a high-fat diet treated with metformin (300 mg kg $^{-1}$ day $^{-1}$) for 11 days. Data are mean \pm s.e.m., $n = 6$ per group except $Gdf15^{+/+}$ vehicle, $n = 7$; P values by two-way ANOVA with Tukey's correction for multiple comparisons. **b**, Cumulative food intake of mice as in **a**, P values by two-way ANOVA with Tukey's correction for multiple comparisons. **c**, Percentage change in body weight of $Gfral^{+/+}$ and $Gfral^{-/-}$ mice on a high-fat diet treated with metformin (300 mg kg $^{-1}$ day $^{-1}$) for 11 days. Data are mean \pm s.e.m., $n = 6$ per group; P values by two-way ANOVA with Tukey's correction for multiple comparisons. **d**, Percentage change in body weight of metformin-treated obese mice dosed with an anti-GFRAL antagonist antibody weekly for five weeks (yellow), starting four weeks after initial metformin exposure (grey). Data are mean \pm s.e.m., vehicle + control IgG and metformin + anti-GFRAL, $n = 7$; other groups, $n = 8$; P values by two-way ANOVA with Tukey's correction for multiple comparisons. Calo, period in which energy expenditure measured (see **e**); arrow, start of oral GTT (Fig. 3e–h). **e**, ANCOVA of energy expenditure against body weight of mice treated as in **d**, $n = 6$ mice per group. Data points show individual mice; P values for metformin calculated using ANCOVA with body weight as a covariate and treatment as a fixed factor.

less pronounced (Extended Data Fig. 2), suggesting an interaction between metformin and the high-fat diet.

To determine the extent to which metformin-induced increase in GDF15 affects body weight, $Gdf15^{+/+}$ and $Gdf15^{-/-}$ mice were switched from chow to a high-fat diet and dosed with metformin for 11 days. The high-fat diet induced similar weight gain in both genotypes (Fig. 2a). Metformin completely prevented weight gain in $Gdf15^{+/+}$ mice, but $Gdf15^{-/-}$ mice were insensitive to the weight-reducing effects of metformin (Fig. 2a, Extended Data Fig. 3a). Metformin significantly reduced cumulative food intake in wild-type mice but this effect was abolished in $Gdf15^{-/-}$ mice (Fig. 2b).

The identical protocol was applied to mice lacking GFRAL, the ligand-binding component of the hindbrain-expressed GDF15 receptor complex. Consistent with the results in mice lacking GDF15, metformin was unable to prevent weight gain in $Gfral^{-/-}$ mice (Fig. 2c, Extended Data Fig. 3b), despite similar levels of serum GDF15 to wild-type mice

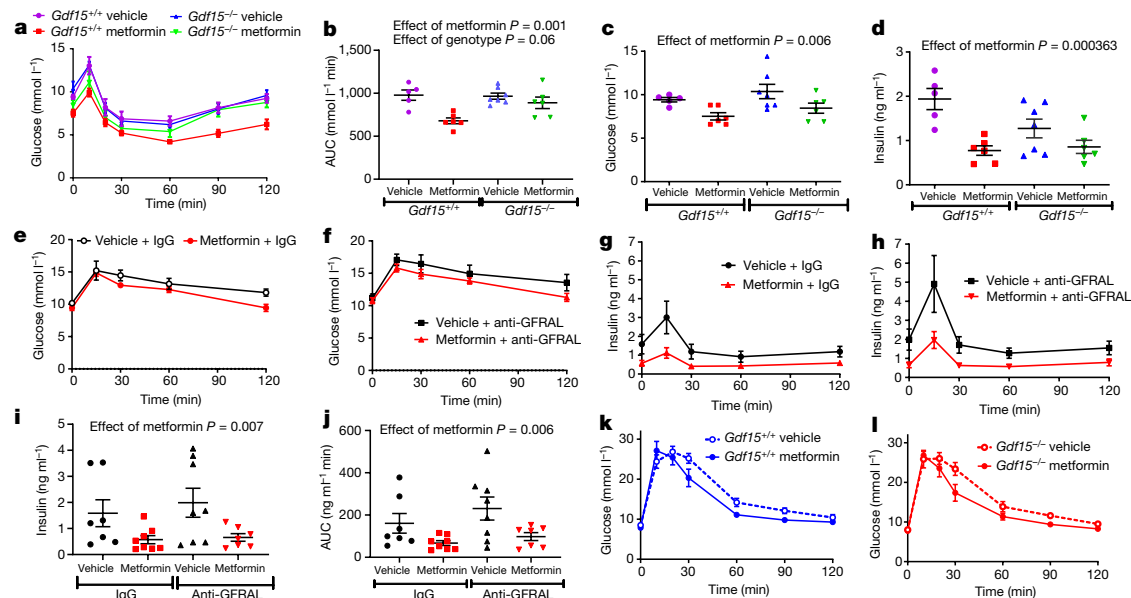


Fig. 3 | Effects of metformin on glucose homeostasis. **a**, ITT (0.5 U kg⁻¹ insulin) after 11 days of metformin treatment (300 mg kg⁻¹) in *Gdf15*^{+/+} and *Gdf15*^{-/-} mice on a high-fat diet. Data are mean ± s.e.m., *n* = 6 per group, except *Gdf15*^{-/-} vehicle, *n* = 7 and *Gdf15*^{+/+} vehicle, *n* = 5. **b**, AUC analysis of glucose over time in mice from **a**. Data are mean ± s.e.m.; *P* values by two-way ANOVA; interaction of genotype and metformin, *P* = 0.037. **c**, Fasting glucose (time 0) of ITT from **a**. Data are mean ± s.e.m.; *P* values by two-way ANOVA, effect of genotype, *P* = 0.144; interaction of genotype and metformin, *P* = 0.988. **d**, Fasting insulin (time 0) in ITT from **a**. Data are mean ± s.e.m.; *P* values by two-way ANOVA, effect of genotype, *P* = 0.131; interaction of genotype and metformin, *P* = 0.056. **e, f**, Glucose over time after oral GTT in metformin-treated obese mice given either IgG (**e**) or anti-GFRAL (**f**) once weekly for five weeks (as in Fig. 2d).

(Extended Data Fig. 4a, b). In this experiment, the reduction in cumulative food intake did not reach statistical significance (Extended Data Fig. 4c).

To investigate the contribution of GDF15–GFRAL signaling to sustained, metformin-dependent weight regulation, we performed a 9-week study in which mice received approximately 250–300 mg kg⁻¹ day⁻¹ of metformin incorporated into their high-fat diet. The mice lost around 9% of their body weight after 1 month on this diet (Fig. 2d Extended Data Fig. 3c). At this time, an anti-GFRAL antagonist antibody or IgG control was administered. Metformin-consuming mice treated with anti-GFRAL regained about 12% of body weight after 5 weeks, whereas the weight loss seen in IgG control treated mice was maintained, reaching approximately 7% below the starting weight (Fig. 2d). The significant reduction in fat mass seen with metformin treatment and control antibody was not seen in the anti-GFRAL group. (Extended Data Fig. 4d). The delivery of metformin in chow resulted in an initial reduction in food intake in all metformin-treated groups, presumably because of a taste effect. This reduction in food intake will have affected metformin levels and probably affected GDF15 levels, with potential to bias the results. However, it is reassuring to note that any persistence of this would have worked against the detection of a specific effect of GFRAL antagonism, which was clearly demonstrable.

We undertook indirect calorimetry in metformin- and placebo-treated mice treated with anti-GFRAL antibody to establish whether there are additional effects on energy expenditure. Data were analysed by analysis of covariance (ANCOVA) with body weight as the covariate. Metformin treatment resulted in a significant increase in metabolic rate, which was blocked by antagonism of GFRAL (Fig. 2e). Thus under conditions in which GDF15 levels are increased by metformin, body

AUC analysis by two-way ANOVA; effect of antibody, *P* = 0.031; effect of metformin, *P* = 0.072; interaction of antibody and metformin, *P* = 0.91.

g, h, Insulin over time after oral GTT in mice treated as in **e, f**. Data are mean ± s.e.m. **i**, Fasting insulin (time 0) after GTT in mice treated as in **e, f**. Data are mean ± s.e.m., *P* values by two-way ANOVA; effect of antibody, *P* = 0.544; interaction of genotype and metformin, *P* = 0.691. **j**, AUC analysis of insulin over time in **g, h**. Data are mean ± s.e.m.; *P* values by two-way ANOVA; effect of antibody, *P* = 0.197; interaction of genotype and metformin, *P* = 0.607. **k, l**, Glucose over time after intraperitoneal GTT in mice on a high-fat diet given a single dose of oral metformin (300 mg kg⁻¹) 6 h before the GTT. Data are mean ± s.e.m., *n* = 8 per group.

weight reduction is contributed to by both reduced food intake and an inappropriately high energy expenditure.

GDF15 and glucose homeostasis

To examine the extent to which the insulin-sensitizing effects of metformin are dependent on GDF15, we repeated the experiment described in Fig. 2a (see Extended Data Fig. 5), measuring insulin tolerance in metformin- and vehicle-treated GDF15-null mice and their wild-type littermates (Fig. 3a). Circulating metformin levels in both genotypes were identical (Extended Data Fig. 5d) and consistent with the high end of the human therapeutic range¹⁴. Metformin significantly increased insulin sensitivity, as assessed by the area under the plasma glucose curve, with no significant effect of genotype (Fig. 3b). Similarly, metformin reduced fasting blood glucose and fasting insulin in a GDF15-independent manner (Fig. 3c, d).

We also performed oral glucose-tolerance tests (GTTs) on metformin-treated mice given either control IgG or anti-GFRAL antibody for five weeks (Figs. 2d, 3e, f, Extended Data Fig. 6a). Although the effect of metformin glucose disposal on the oral GTT as assessed by the area under the plasma glucose curve did not reach statistical significance (two-way ANOVA, *P* = 0.072), there was a significant effect of metformin on insulin (both fasting level and area under the curve (AUC)) after glucose bolus, that was independent of anti GFRAL antibody (Fig. 3g–j).

As these mice had different body weights at the time of assessment (Fig. 2d, Extended Data Fig. 3c), we performed intraperitoneal GTTs in a cohort of weight-matched *Gdf15*^{+/+} and *Gdf15*^{-/-} mice that had been fed a high-fat diet for two weeks before receiving a single dose of metformin (300 mg kg⁻¹) (Fig. 3k, l, Extended Data Fig. 6b–d). In these mice, there

was a significant effect of metformin on glucose levels (plasma glucose AUC) that was independent of GDF15 (Extended Data Fig. 6e).

The effect of metformin in decreasing fasting glucose and insulin and improving glucose tolerance do not require GDF15. Given the a priori expected effect of weight loss on insulin sensitivity it is noteworthy that the effect of GDF15 status on insulin sensitivity as measured by insulin-tolerance test (ITT) (Fig. 3b) fell just short of statistical significance. In the follow up of the Diabetes Prevention Program study in non-diabetic individuals, weight loss after 5 years of metformin therapy was approximately 6.5% of baseline weight⁵. We therefore estimated the effect of a 6.5% weight loss on improvements in fasting insulin over 5 years in the Ely Study, a prospective observational population-based cohort study of men ($n = 465$) and women ($n = 634$) in the UK (mean age 52 years, mean body mass index 26 at baseline)¹⁵, showing that this magnitude of weight loss was associated with a reduction in fasting plasma insulin of -5.74 ($-9.03, -2.45$) pmol l^{-1} (mean \pm 95% confidence interval (CI)) in women and -8.78 ($-16.24, -1.33$) pmol l^{-1} in men. We conclude that although there are GDF15-independent effects of metformin on circulating levels of glucose and insulin, GDF15-dependent weight loss probably contributes to enhancing insulin sensitivity.

Source of GDF15 production

We examined *Gdf15* gene expression in a tissue panel obtained from mice fed a high-fat diet (for four weeks) and euthanized 6 h after a single-gavage dose of metformin (600 mg kg^{-1}). Circulating concentrations of GDF15 increased about 5.5-fold compared with vehicle-treated mice (Extended Data Fig. 6f) and *Gdf15* mRNA was significantly increased by metformin in small intestine, colon and kidney (Fig. 4a). In situ hybridization studies demonstrated strong *Gdf15* expression in crypt enterocytes in the colon and small intestine and in periglomerular renal tubular cells (Fig. 4b, Extended Data Fig. 7a, b). We confirmed these sites of tissue expression in mice fed a high-fat diet (those used in Fig. 2a) and treated with metformin for 11 days (Extended Data Fig. 8). Further, in organoids derived from human (Fig. 4c) and mouse (Fig. 4d) intestine, grown in two-dimensional (2D) transwells and treated with metformin, we observed a significant induction of *Gdf15* mRNA expression and GDF15 protein secretion.

Given the proposed importance of the liver for the metabolic action of metformin, it was notable that the dominant GDF15 expression signal was not from the liver (Fig. 4a, Extended Data Figs. 7a, 8). To determine whether hepatocytes are capable of responding to biguanide drugs with an increase in GDF15, we incubated freshly isolated mouse hepatocytes (Extended Data Fig. 9a) and stem-cell derived human hepatocytes (Extended Data Fig. 9b) with metformin and found a clear induction of GDF15 expression. Additionally, acute administration of the more cell-penetrant biguanide drug phenformin to mice increased circulating GDF15 levels (Extended Data Fig. 9c) and markedly increased *Gdf15* mRNA expression in hepatocytes (Extended Data Fig. 9d, e). We conclude that biguanides can induce GDF15 expression in many cell types but, at least when given orally to mice, *Gdf15* mRNA is mostly induced in the distal small intestine, colon and kidney.

GDF15 expression has been reported to be a downstream target of the cellular integrated stress response (ISR) pathway^{16–18}. *Gdf15* mRNA levels were increased in kidney and colon 24 h after a single oral dose of metformin and these changes correlated positively with the fold elevation of *Chop* (also known as *Ddit3*) mRNA (Extended Data Fig. 10a, b). As phenformin has broader cell permeability than metformin¹⁹, we used it to explore the effects of biguanides on the ISR and its relationship to GDF15 expression in cells. In mouse embryonic fibroblasts (MEFs), which do not express the organic cation transporters needed for the uptake of metformin, phenformin (but not metformin) increased EIF2 α phosphorylation, ATF4 and CHOP expression (Extended Data Fig. 10c) and *Gdf15* mRNA (Extended Data Fig. 10d), though the changes in EIF2 α phosphorylation and ATF4 and CHOP expression were modest

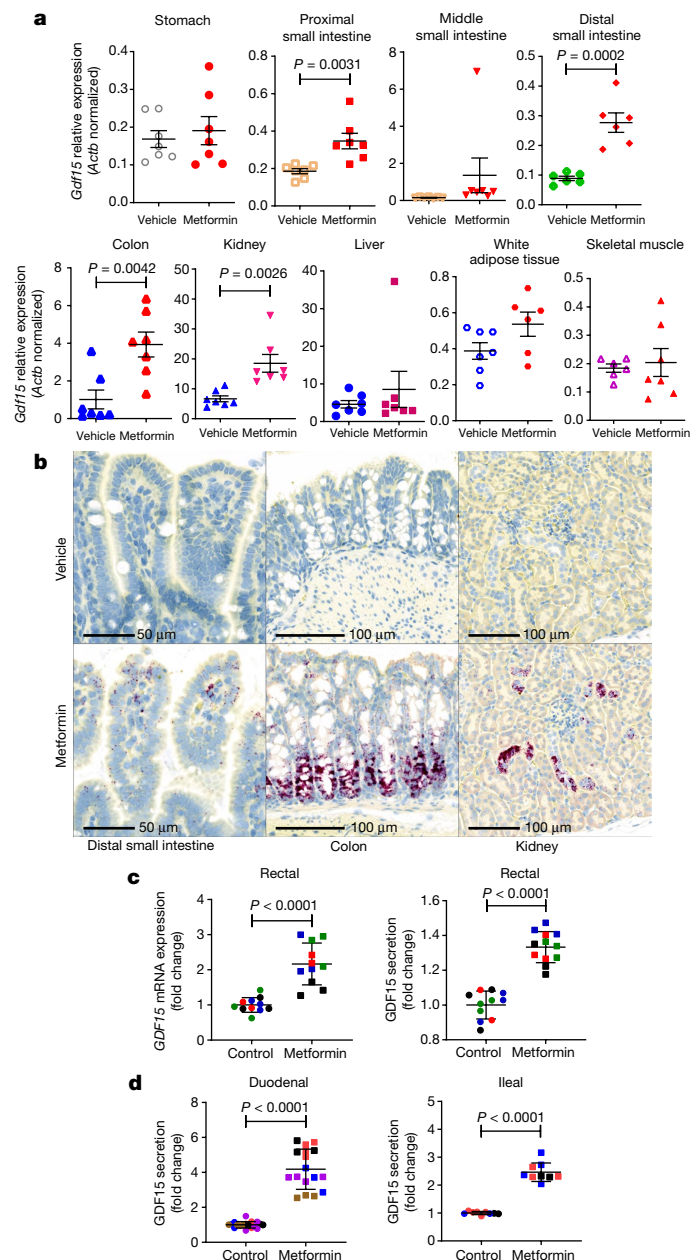


Fig. 4 | Metformin increases GDF15 expression in the enterocytes of distal intestine and in renal tubular epithelial cells. a, *Gdf15* mRNA expression (normalized to expression levels of *Actb*) 6 h after a single dose of oral metformin (600 mg kg^{-1}) in tissues from wild-type mice on a high-fat diet. Data are mean \pm s.e.m., $n = 7$ per group; P value (95% CI) by two-tailed t -test. **b**, In situ hybridization for *Gdf15* mRNA (red spots). $n = 7$ per group. Representative images from the mouse with circulating GDF15 level closest to group median, treated with vehicle or metformin. Mice are from groups described in **a**. **c**, *Gdf15* mRNA expression (left) and GDF15 protein in supernatant (right) of human-derived 2D monolayer rectal organoids treated with metformin. Each colour represents an independent experiment. Data are mean \pm s.d., $n = 4$; P values (95% CI) by two-tailed t -test. **d**, GDF15 protein in supernatants of mouse-derived 2D monolayer duodenal (left) and ileal (right) organoids treated with metformin. Each colour represents an independent experiment. Data are mean \pm s.d., duodenal, $n = 5$; ileal, $n = 3$; P values (95% CI) by two-tailed t -test.

compared with those induced by tunicamycin despite similar levels of *Gdf15* mRNA induction. Both genetic deletion of *Atf4* and small interfering (si)RNA-mediated knockdown of *Chop* significantly reduced phenformin-mediated induction of *Gdf15* mRNA expression (Extended Data Fig. 10e, f). In addition, phenformin induction of GDF15

was markedly reduced by co-treatment with the EIF2 α inhibitor ISRIB but, notably, not by the PERK inhibitor GSK2606414 (Extended Data Fig. 10g). Further, GDF15 secretion in response to metformin in mouse duodenal organoids was also significantly reduced by co-treatment with ISRIB (Extended Data Fig. 10h). However, gut organoids derived from CHOP-null mice are still able to increase GDF15 secretion in response to metformin (Extended Data Fig. 10i) indicating the existence of CHOP-independent pathways under some circumstances. These data suggest that the effects of biguanides on GDF15 expression are at least partly dependent on the ISR pathway but are independent of PERK. However, the relative importance of components of the ISR pathway may vary depending on specific cell type, dose and agent used.

Our observations represent an advance in the understanding of the action of metformin, one of the world's most frequently prescribed drugs. Metformin increases circulating GLP1 levels^{20–22}, but its metabolic effects in mice are unimpaired in mice lacking the GLP1 receptor²³. Metformin alters the intestinal microbiome^{24,25} but it is challenging to firmly establish a causal relationship between this effect and the beneficial effects of the drug²⁶.

In this study, we present a body of data from humans, cells, organoids and mice that securely establish a major role for GDF15 in the mediation of the beneficial effects of metformin on energy balance. Whereas these effects probably contribute to the role of metformin as an insulin sensitizer, it has other effects in decreasing glucose and insulin in the absence of GDF15.

Whereas many mechanisms have been suggested for the glucoregulatory mechanisms of metformin²⁷, there has been less attention paid to its effects on weight. Our discoveries relating to metformin's effects via GDF15 provide a compelling explanation for this important aspect of its action.

It is notable that the lower small intestine and colon are a major site of metformin-induced GDF15 expression. An emerging body of work strongly implicates the intestine as a major site of metformin action. Metformin increased glucose uptake into colonic epithelium from the circulation²⁸ and a gut-restricted formulation of metformin had greater glucose-lowering efficacy than systemically absorbed formulations²⁹. Our finding that the intestine is a major site of metformin-induced GDF15 expression provides a further mechanism through which metformin's action on the intestinal epithelium may mediate some of its benefits.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1911-y>.

- Knowler, W. C. et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N. Engl. J. Med.* **346**, 393–403 (2002).
- Ramachandran, A. et al. The Indian Diabetes Prevention Programme shows that lifestyle modification and metformin prevent type 2 diabetes in Asian Indian subjects with impaired glucose tolerance (IDPP-1). *Diabetologia* **49**, 289–297 (2006).

- Lachin, J. M. et al. Factors associated with diabetes onset during metformin versus placebo therapy in the diabetes prevention program. *Diabetes* **56**, 1153–1159 (2007).
- Rena, G., Hardie, D. G. & Pearson, E. R. The mechanisms of action of metformin. *Diabetologia* **60**, 1577–1585 (2017).
- Apolzan, J. W. et al. Long-term weight loss with metformin or lifestyle intervention in the Diabetes Prevention Program Outcomes Study. *Ann. Intern. Med.* **170**, 682–690 (2019).
- Gerstein, H. C. et al. Growth differentiation factor 15 as a novel biomarker for metformin. *Diabetes Care* **40**, 280–283 (2017).
- Tsai, V. W. W., Husaini, Y., Sainsbury, A., Brown, D. A. & Breit, S. N. The MIC-1/GDF15–GFRAL pathway in energy homeostasis: implications for obesity, cachexia, and other associated diseases. *Cell Metab.* **28**, 353–368 (2018).
- Mullican, S. E. et al. GFRAL is the receptor for GDF15 and the ligand promotes weight loss in mice and nonhuman primates. *Nat. Med.* **23**, 1150–1157 (2017).
- Emmerson, P. J. et al. The metabolic effects of GDF15 are mediated by the orphan receptor GFRAL. *Nat. Med.* **23**, 1215–1219 (2017).
- Yang, L. et al. GFRAL is the receptor for GDF15 and is required for the anti-obesity effects of the ligand. *Nat. Med.* **23**, 1158–1166 (2017).
- Hsu, J. Y. et al. Non-homeostatic body weight regulation through a brainstem-restricted receptor for GDF15. *Nature* **550**, 255–259 (2017).
- Konopka, A. R. et al. Hyperglucagonemia mitigates the effect of metformin on glucose production in prediabetes. *Cell Rep.* **15**, 1394–1400 (2016).
- Preiss, D. et al. Metformin for non-diabetic patients with coronary heart disease (the CAMERA study): a randomised controlled trial. *Lancet Diabetes Endocrinol.* **2**, 116–124 (2014).
- McCreight, L. J. et al. Pharmacokinetics of metformin in patients with gastrointestinal intolerance. *Diabetes Obes. Metab.* **20**, 1593–1601 (2018).
- Forouhi, N. G., Luan, J., Hennings, S. & Wareham, N. J. Incidence of Type 2 diabetes in England and its association with baseline impaired fasting glucose: the Ely study 1990–2000. *Diabet. Med.* **24**, 200–207 (2007).
- Chung, H. K. et al. Growth differentiation factor 15 is a myomitokine governing systemic energy homeostasis. *J. Cell Biol.* **216**, 149–165 (2017).
- Li, D., Zhang, H. & Zhong, Y. Hepatic GDF15 is regulated by CHOP of the unfolded protein response and alleviates NAFLD progression in obese mice. *Biochem. Biophys. Res. Commun.* **498**, 388–394 (2018).
- Patel, S. et al. GDF15 provides an endocrine signal of nutritional stress in mice and humans. *Cell Metab.* **29**, 707–718 (2019).
- Shu, Y. et al. Effect of genetic variation in the organic cation transporter 1 (OCT1) on metformin action. *J. Clin. Invest.* **117**, 1422–1431 (2007).
- DeFronzo, R. A. et al. Once-daily delayed-release metformin lowers plasma glucose and enhances fasting and postprandial GLP-1 and PYY: results from two randomised trials. *Diabetologia* **59**, 1645–1654 (2016).
- Preiss, D. et al. Sustained influence of metformin therapy on circulating glucagon-like peptide-1 levels in individuals with and without type 2 diabetes. *Diabetes Obes. Metab.* **19**, 356–363 (2017).
- Bahne, E. et al. Metformin-induced glucagon-like peptide-1 secretion contributes to the actions of metformin in type 2 diabetes. *JCI Insight* **3**, 93936 (2018).
- Maida, A., Lamont, B. J., Cao, X. & Drucker, D. J. Metformin regulates the incretin receptor axis via a pathway dependent on peroxisome proliferator-activated receptor- α in mice. *Diabetologia* **54**, 339–349 (2011).
- de la Cuesta-Zuluaga, J. et al. Metformin is associated with higher relative abundance of mucin-degrading *Akkermansia muciniphila* and several short-chain fatty acid-producing microbiota in the gut. *Diabetes Care* **40**, 54–62 (2017).
- Shin, N. R. et al. An increase in the *Akkermansia* spp. population induced by metformin treatment improves glucose homeostasis in diet-induced obese mice. *Gut* **63**, 727–735 (2014).
- Forslund, K. et al. Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature* **528**, 262–266 (2015).
- Foretz, M., Guigas, B. & Viollet, B. Understanding the glucoregulatory mechanisms of metformin in type 2 diabetes mellitus. *Nat. Rev. Endocrinol.* **15**, 569–589 (2019).
- Massollo, M. et al. Metformin temporal and localized effects on gut glucose metabolism assessed using ¹⁸F-FDG PET in mice. *J. Nucl. Med.* **54**, 259–266 (2013).
- Buse, J. B. et al. The primary glucose-lowering effect of metformin resides in the gut, not the circulation: results from short-term pharmacokinetic and 12-week dose-ranging studies. *Diabetes Care* **39**, 198–205 (2016).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

Methods

Human studies

We analysed samples from nine participants from a study with a placebo-controlled, double-blind crossover design (previously described in ref. ¹²). In brief, placebo or metformin (week 1, 500 mg twice daily; week 2, 1,000 mg twice daily) was administered following a six-week period of washout. Samples were collected in the morning after overnight fasting. The study was approved by the Mayo Clinic Institutional Review Board and all participants provided written, informed consent (NCT01956929).

CAMERA was a randomized, double-blinded, placebo-controlled trial designed to investigate the effect of metformin on surrogate markers of cardiovascular disease in patients without diabetes, aged 35 to 75, with established coronary heart disease and a large waist circumference (≥ 94 cm in men, ≥ 80 cm in women) (NCT00723307). This single-centre trial enrolled 173 adults who were followed up for 18 months each. A detailed description of the trial and its results has been published previously¹³. In brief, participants were randomized 1:1 to 850 mg metformin or matched placebo twice daily with meals. Participants attended six monthly visits after overnight fasts and before taking their morning dose of metformin. Blood samples collected during the trial were centrifuged at 4 °C soon after sampling, separated and stored at -80 °C.

All participants provided written informed consent. The study was approved by the Medicines and Healthcare Products Regulatory Agency and West Glasgow Research Ethics Committee, and done in accordance with the principles of the Declaration of Helsinki and good clinical practice guidelines.

Serum GDF15 assays were measured in a fashion blinded to treatment allocation or timing of samples by the Cambridge Biochemical Assay Laboratory, University of Cambridge. Measurements were performed with antibodies and standards from R&D Systems (R&D Systems) using a microtitre-plate-based two-site electrochemiluminescence immunoassay using the MesoScale Discovery assay platform (MSD).

Mouse studies

Studies were carried out at two sites: NGM Biopharmaceuticals, California, and the University of Cambridge.

At NGM, all experiments were conducted with NGM IACUC approved protocols and all relevant ethical regulations were complied with throughout the course of the studies, including efforts to reduce the number of animals used. Experimental animals were kept under controlled light (12 h:12 h light:dark cycle, dark 18:30–06:30), temperature (22 ± 3 °C) and humidity ($50 \pm 20\%$) conditions. They were fed ad libitum on 2018 Teklad Global 18% Protein Rodent Diet containing 24 kcal% fat, 18 kcal% protein and 58 kcal% carbohydrate, or on high-fat rodent diet containing 60 kcal% fat, 20 kcal% protein and 20 kcal% carbohydrates from Research Diets D12492i, hereafter referred to as 60% HFD.

In Cambridge, all mouse studies were performed in accordance with UK Home Office Legislation regulated under the Animals (Scientific Procedures) Act 1986 Amendment, Regulations 2012, following ethical review by the University of Cambridge Animal Welfare and Ethical Review Body (AWERB). They were maintained in a 12 h:12 h light:dark cycle (lights on 07:00–19:00), temperature-controlled (22 °C) facility, with ad libitum access to food (RM3(E) Expanded Chow (Special Diets Services)) and water. Any mice bought from an outside supplier were acclimatised in a holding room for at least one week before study. During study periods they were fed ad libitum high-fat diet, either D12451i (45 kcal% fat, 20 kcal% protein and 35 kcal% carbohydrates, herein referred to as 45% HFD) or D12492i (Research Diets) as highlighted in the individual study.

Sample sizes were determined on the basis of homogeneity and consistency of characteristics in the selected models and were sufficient to detect statistically significant differences in body weight,

food intake and serum parameters between groups. Experiments were performed with animals of a single gender in each study. Animals were randomized into the treatment groups on the basis of body weight such that the mean body weights of each group were as close to each other as possible, but without using an excess number of animals. No samples or animals were excluded from analyses. Researchers were not blinded to group allocations.

Mouse study 1, acute two-dose metformin and high-fat diet

Male C57BL6/J mice fed 60% HFD for 17 weeks were studied aged 23 weeks (body weight, mean \pm s.e.m., 45.6 ± 0.8 g). Metformin (Sigma-Aldrich no. 1396309) was reconstituted in water at 30 mg ml⁻¹ for oral gavage and given in the early part of the light cycle. Terminal blood was collected by cardiac puncture into EDTA-coated tubes. GDF15 levels were measured using Mouse/Rat GDF15 Quantikine ELISA Kit (no. MGD-150, R&D Systems) according to the manufacturers' instructions. RNA was isolated from tissues using the Qiagen RNeasy Kit. RNA was quantified and 500 ng was used for cDNA synthesis (SuperScript VILO; 11754050, ThermoFisher) followed by quantitative (q)PCR. All Taqman probes were purchased from Applied Biosystems. All genes are expressed relative to 18S control probe and were run in triplicate.

Mouse study 2, acute metformin and normal diet

Ad libitum group. Male C57BL6/J mice (Charles River) were studied at 11 weeks old. Five-hundred milligrams of metformin was dissolved in 20 ml water to make a working stock of 25 mg ml⁻¹. One hour after onset of light cycle, mice received a single dose by oral gavage of either metformin at 300 mg kg⁻¹ (Sigma, PHR1084-500MG) or a matched volume of vehicle (water). Weight (mean \pm s.e.m.) of control and treatment groups were 27.2 ± 0.3 g and 26.7 ± 0.2 g, respectively, on the day of study. After gavage, mice were returned to an individual cage and were euthanized at the relevant time point by terminal anaesthesia (Euthatal by intraperitoneal injection). Blood was collected into a Sarstedt Serum Gel 1.1 ml Micro Tube, left for 30 min at room temperature, then spun for 5 min at 10,000g at 40 °C before being frozen and stored at -80 °C until assayed. Mouse GDF15 levels were measured using a Mouse GDF15 DuoSet ELISA (R&D Systems) which had been modified to run as an electrochemiluminescence assay on the Meso Scale Discovery assay platform.

Fasted group. Mice, conditions and methods as in the ad libitum group, except that male mice were studied at 9 weeks old and 12 h before administration of metformin; mice and bedding were transferred to new cages with no food in the hopper. Weight (mean \pm s.e.m.) after fasting and on day of gavage were 22.3 ± 0.5 g and 23.2 ± 0.7 g for control and treatment groups, respectively.

Mouse study 3, metformin and high-fat diet, *Gdf15*^{-/-} and wild-type mice

C57BL/6N-Gdf15tm1a(KOMP)Wtsi/H mice (referred to as *Gdf15*^{-/-} mice) were obtained from the MRC Harwell Institute, which distributes these mice on behalf of the European Mouse Mutant Archive (<https://www.infrafrontier.eu/>). The MRC Harwell Institute is also a member of the International Mouse Phenotyping Consortium (IMPC) and has received funding from the MRC for generating and/or phenotyping the C57BL/6N-Gdf15tm1a(KOMP)Wtsi/H mice. The research reported in this publication is solely the responsibility of the authors and does not necessarily represent the official views of the Medical Research Council. Associated primary phenotypic information may be found at <https://www.mousephenotype.org/>. Details of the alleles have been published^{30–32}.

Experimental cohorts of male *Gdf15*^{-/-} and wild-type mice were generated by het \times het breeding pairs. Mice were aged between 4.5 and 6.5 months. One week before study start, mice were single-housed and three days before the first dose of metformin treatment, mice were

Article

transferred from standard chow to 60% high-fat diet. On the day of first gavage, body weight of study groups (mean \pm s.e.m.) were 38.2 ± 1.0 g vs 38.8 ± 0.6 g for wild-type vehicle and metformin treatments, respectively, and 37.9 ± 0.8 g vs 37.0 ± 1.4 g for *Gdf15*^{+/+} vehicle and metformin treatments, respectively. Each mouse received a daily gavage of either vehicle or metformin for 11 days, and their body weight and food intake was measured daily in the early part of the light cycle. One data point of 25 food intake points collected on day 11 of the study was lost owing to technical error (*Gdf15*^{+/+} mouse, metformin). On day 11, mice were euthanized by terminal anaesthesia 4 h after gavage, and blood was obtained as in mouse study 2. Tissues were fresh frozen on dry ice and kept at -80°C until the day of RNA extraction.

Mouse study 4, metformin and high-fat diet, *Gfrr1*^{-/-} mice

Gfrr1^{-/-} mice on a mixed 129/SvEv-C57BL/6 background were purchased from Taconic (TF3754) and backcrossed for 10 generations to >99% C57BL/6 background at the NGM animal facility. Experimental cohorts were generated by het \times het breeding pairs. Study design was as in study 3, except that terminal blood was collected into EDTA-coated tubes.

Mouse study 5, anti GFRAL antibody, metformin and high-fat diet

Anti-GFRAL antibody generation. Anti-GFRAL monoclonal antibodies were generated by immunizing C57BL/6 mice with recombinant purified GFRAL ECD-hFc fusion protein, which was purified via sequential protein-A affinity and size exclusion chromatography techniques using MabSelect SuRe and Superdex 200 purification media, respectively (GE Healthcare), as described in patent US10174119B2 (<https://patents.google.com/patent/US10174119B2/en>). An in-house pTT5 hlgK hlgG1 expression vector was engineered to include the DEVDG (caspase-3) proteolytic site N-terminal to the Fc domain. The heavy chains of anti-GFRAL monoclonal antibodies were subcloned via EcoRI and HindIII sites of an in-house engineered pTT5 hlgK hlgG1 caspase-cleavable vector. Light chains of anti-GFRAL monoclonal antibodies were also subcloned in the EcoRI and HindIII sites in the pTT5 hlgK hKappa vector. The antibodies were transiently expressed in Expi293 cells (Thermo Fisher Scientific) transfected with the pTT5 expression vector, and purified from conditioned media by sequential protein-A affinity and size-exclusion chromatography using MabSelect SuRe and Superdex 200 purification media, respectively (GE Healthcare). All purified antibody material was verified endotoxin-free and formulated in PBS for in vitro and in vivo studies. Characterization of anti-GFRAL functional blocking antibodies was carried out using a cell-based RET/GFRAL luciferase gene reporter assays, in vitro binding studies (ELISA and Biacore) and in vivo studies, as described in patent number US10174119B2 (<https://patents.google.com/patent/US10174119B2/en>).

In all studies with anti-GFRAL, purified recombinant non-targeting IgG on the same antibody framework was used as control. Metformin was mixed with food paste made from the 60 kcal% fat diet (Research Diet no. D12492) using a food blender at a concentration to achieve an approximate consumption of 300 mg kg⁻¹ metformin per day per mouse. Male mice were single-housed throughout and at the start of study period, body weight (mean \pm s.e.m.) was 43.7 ± 1.4 g (vehicle + control IgG), 42.3 ± 1.4 g (vehicle + anti-GFRAL), 41.9 ± 1.1 g (metformin + control IgG) and 43.3 ± 1.3 g (metformin + anti-GFRAL). Recombinant antibodies were administered by subcutaneous injection in the early part of the light cycle. Body composition (lean and fat mass) was analysed by ECHO MRI M113 mouse system (Echo Medical Systems). The metabolic parameters oxygen consumption (VO₂) and carbon dioxide production (VCO₂) were measured by an indirect calorimetry system (LabMaster TSE System) in open-circuit sealed chambers. Measurements were performed for the dark (from 18:00 to 06:00) or light (from 06:00 to 18:00) period under ad libitum feeding conditions. Mice were placed in individual metabolic cages and allowed to acclimate for a period of 24 h before data collection every 30 min.

Finally, mice underwent a GTT. Mice were fasted for 6 h (07:00–13:00) in a clean cage. Blood samples (~ 30 μl) were collected as baseline before oral GTT. Mice were orally gavaged with 1 g kg⁻¹ of 20% glucose solution with a dosing volume of 5 ml kg⁻¹. Blood samples were then collected through a tail nick into K₂-EDTA-coated tubes (SARSTEDT Microvette; no. 20.1278.100) at 15, 30, 60 and 120 min after glucose challenge. Blood samples were centrifuged at 4°C and the separated plasma was stored at -20°C until used for plasma glucose and insulin assays. Glucose assay reagents were obtained from Wako (no. 439-90901) and the insulin ELISA kit was obtained from ALPCO (no. 80-INSMSU-E01).

Mouse study 6, ITT after metformin and high-fat diet, *Gdf15*^{-/-} and wild-type mice

Mice generation and protocol are the same as in study 3, except mice were aged 4 to 6 months. On the day of the first gavage, body weights (mean \pm s.e.m.) of study groups were 35.1 ± 1.2 g (wild-type, vehicle), 35.05 ± 1.2 g (wild-type, metformin), 35.08 ± 1.02 g (*Gdf15*^{-/-}, vehicle) and 35.02 ± 1.47 g (*Gdf15*^{-/-}, metformin). On day 11, after the final dose of metformin, mice were fasted for 4 h. Baseline venous blood sample was collected into heparinised capillary tube for insulin measurement and blood glucose was measured using approximately 2 μl blood drops using a glucometer (AlphaTrak2; Abbot Laboratories) and glucose strips (AlphaTrak2 test 2 strips, Abbot Laboratories, Zoetis). Mice were given intraperitoneal injection of insulin (0.5 U kg⁻¹, Actrapid, NovoNordisk) and serial mouse glucose levels were measured at time points indicated. Mice were killed by terminal anaesthesia as in study 2. Mouse insulin was measured using a Meso Scale Discovery two-plex mouse metabolic immunoassay kit according to the manufacturer's instructions and using calibrators provided by Meso Scale Diagnostics. Serum metformin levels were quantified using a stable isotope dilution liquid chromatography–mass spectrometry (LC–MS/MS) method described previously³³.

Mouse study 7, GTT after metformin and high-fat diet, *Gdf15*^{-/-} and wild-type mice

Mice generation was as in study 3, except using female mice aged 3.5 to 5.5 months. Two groups of mice (*Gdf15*^{+/+} and *Gdf15*^{-/-} littermates, body weight (mean \pm s.e.m.) 24.1 ± 1.4 g and 24.3 ± 1.3 g, respectively) were fed 60% HFD for two weeks. Each genotype was then further split into vehicle or metformin (300 mg kg⁻¹) treatment groups, given a single gavage dose at 08:00 and fasted for 6 h. At time of GTT, body weights (mean \pm s.e.m.) of study groups were 26.4 ± 1.5 g (wild-type, vehicle), 26.5 ± 1.0 g (wild-type, metformin), 25.6 ± 1.2 g (*Gdf15*^{-/-}, vehicle) and 27.1 ± 1.3 g (*Gdf15*^{-/-}, metformin); one-way ANOVA, $P = 0.8722$. Baseline testing was as in mouse study 6. Mice then received a single dose of 20% glucose intraperitoneally (2 mg g⁻¹) with serial measurement of glucose levels at time points indicated. Euthanasia and insulin analysis were performed as in mouse study 6.

Mouse study 8, acute single high-dose metformin and high-fat diet

Male C57BL6/J mice (Charles River) aged 14 weeks were switched from standard chow to 45% HFD (D12451i) for 1 week then 60% HFD (D12492i) for 3 weeks. At the time of the study (18 weeks old) body weights (mean \pm s.e.m.) were 40.4 ± 1.2 g and 41.1 ± 1.3 g for the vehicle and metformin groups, respectively. Five-hundred milligrams of metformin (Sigma, PHR1084-500MG) were dissolved in 8.35 ml water to make a working stock of 60 mg ml⁻¹. Mice received a single dose by oral gavage of either 600 mg kg⁻¹ metformin or a matched volume of vehicle (water). They were returned to ad libitum 60% HFD and 6 h later blood was collected as in study 2. Tissue samples for RNA analysis were collected into Lysing Matrix D homogenization tubes (MP Biomedicals) on dry ice and stored at -80°C until they were processed. Intestine between pylorus of stomach and caecum was laid out into three equal parts, with tissue taken from the midpoint of each third labelled as

'proximal', 'middle' and 'distal' (adapted from ref. ³⁴). The colon section was from the midpoint between caecum and anus. Tissues for in situ hybridization were dissected and placed into 10% formalin/PBS for 24 h at room temperature, transferred to 70% ethanol and processed into paraffin. Five-micrometre sections were cut and mounted onto Superfrost Plus (Thermo Fisher Scientific). Detection of mouse *Gdf15* was performed on formalin-fixed paraffin-embedded sections using Advanced Cell Diagnostics (ACD) RNAscope 2.5 LS Reagent Kit-RED (no. 322150) and RNAscope LS 2.5 Probe Mm-Gdf15-O1 (no. 442948) (ACD). In brief, sections were baked for 1 h at 60 °C before loading onto a Bond RX instrument (Leica Biosystems). Slides were deparaffinized and rehydrated on board before pre-treatments using Epitope Retrieval Solution 2 (no. AR9640, Leica Biosystems) at 95 °C for 15 min, and ACD Enzyme from the LS Reagent kit at 40 °C for 15 min. Probe hybridization and signal amplification was performed according to the manufacturer's instructions. Fast red detection of mouse *Gdf15* was performed on the Bond RX using the Bond Polymer Refine Red Detection Kit (Leica Biosystems, no. DS9390) according to the ACD protocol. Slides were then counterstained with haematoxylin, removed from the Bond RX and were heated at 60 °C for 1 h, dipped in xylene and mounted using EcoMount Mounting Medium (Biacare Medical, no. EM897L).

Slides imaged on an automated slide-scanning microscope (Axioscan Z1 and Hamamatsu orca flash 4.0 V3 camera) using a 20× objective with a numerical aperture of 0.8. Hybridization specificity was confirmed by the absence of staining in *Gdf15*^{-/-} mice.

RNA extraction was carried out with approximately 100 mg of tissue in 1 ml Qiazol Lysis Reagent (Qiagen 793061) using Lysing Matrix D homogenization tube and Fastprep 24 Homogenizer (MP Biomedicals) and Qiagen RNeasy Mini Kit (no. 74106) with DNaseI treatment following manufacturers' protocols. Five-hundred nanograms of RNA was used to generate cDNA using Promega M-MLV reverse transcriptase followed by TaqMan qPCR in triplicate for *GDF15*. Samples were normalized to *Actb*. TaqMan Probes: Mm00442228 m1 GDF15, Mm02619580_g1 Act B, TaqMan; 2× universal PCR Master mix (Applied Biosystems Thermo Fisher, 4318157); QuantStudio 7 Flex Real time PCR system (Applied Biosystems Life Technologies).

Mouse study 9, acute phenformin and normal diet, wild-type mice

Male C57BL6/J mice aged 14 weeks with supplier, protocol and methods as in study 2, except that phenformin (Sigma PHR1573, 500 mg) was used instead of metformin.

Organoid studies

Duodenal and ileal mouse organoid line generation, maintenance and 2D culture was performed as previously described³⁵. *CHOP*-null mice were a gift from J. Goodall, with a line from Jackson Laboratory (B6.129S(Cg)-Ddit3tm2.1Dron/J, stock no. 005530). Human rectal organoids (experiments approved by the Research Ethics Committee under licence number 09/H0308/24) were generated from fresh surgical specimens (Tissue Bank, Addenbrooke's Hospital (Cambridge, UK)) following a modified protocol^{35,36}. In brief, rectal tissue was chopped into 5-mm fragments and incubated in 30 mM EDTA for 3 × 10 min, with tissue shaken in PBS after each EDTA treatment to release intestinal crypts. The isolated crypts were then further digested using TrypLE (Life Technologies) for 5 min at 37 °C to generate small cell clusters. These were then seeded into basement membrane extract (BME, R&D Technology), with 20-μl domes polymerized in multiwell (48) dishes for 30–60 min at 37 °C. Organoid medium³⁶ was then overlaid and changed three times per week. Human organoids were passaged every 14–21 days using TrypLE digestion for 15 min at 37 °C, followed by mechanical shearing with rigorous pipetting to break organoids up into small clusters, which were then seeded as before in BME. For Transwell experiments, TrypLE-digested organoids were seeded onto Matrigel (Corning)-coated (2% for 60 min at 37 °C) polyethylene terephthalate

cell culture inserts, pore size 0.3 μm (Falcon) in organoid medium supplemented with Y-27632 (R&D Technology). Organoids were observed through the transparent cell inserts to ensure 2D culture formation (allowing apical cell access for drug treatments). Medium was changed after 2 days and then switched on day 3 to a differentiation medium with Wnt3A-conditioned medium reduced to 10% and SB202190 or nicotinamide omitted from culture for 5 days.

For GDF15 secretion experiments, 2D cultured organoid cells were treated for 24 h with indicated drugs, with medium then collected and GDF15 measured at the Core Biochemical Assay Laboratory (Cambridge) using the human or mouse GDF15 assay kit as outlined in the CAMERA human study and mouse study 2 above.

RNA was extracted using TRI reagent (Sigma), with any contaminated DNA eliminated using DNA-free removal kit (Invitrogen). Purified RNA was then reverse-transcribed using superscript II (Invitrogen) as per the manufacturer's protocol. Quantitative PCR with reverse transcription was performed on a QuantStudio 7 (Applied Biosystems) using Fast Taqman mastermix and the following probes (Applied Biosystems): human *GDF15* (Hs00171132_m1) and human *ACTB* (Hs01060665_g1). Gene expression was measured relative to β-actin in the same sample using the ΔC_t method, with fold (relative to control) shown for each experiment.

Hepatocyte studies

Primary mouse hepatocyte isolation and culture. Hepatocytes from 8 to 12-week-old C57B6J male mice were isolated by retrograde, non-recirculating in situ collagenase liver perfusion. In brief, livers were perfused with modified Hanks medium without calcium (8.0 g l⁻¹ NaCl, 0.4 g l⁻¹ KCl, 0.2 g l⁻¹ MgSO₄·7H₂O, 0.12 g l⁻¹ Na₂HPO₄·2H₂O, 0.12 g l⁻¹ KH₂PO₄, 3 g l⁻¹ HEPES, 0.342 g l⁻¹ EGTA and 0.05 g l⁻¹) followed by digestion with perfusion medium supplemented with calcium (0.585 g l⁻¹ CaCl₂·2H₂O) and 0.5 mg ml⁻¹ collagenase IV (Sigma, C5138). The digested liver was removed and washed using chilled DMEM:F12 (Sigma) medium containing 2 mM L-glutamine, 10% FBS, 1% penicillin/streptomycin (Invitrogen). Viable cells were collected by Percoll (Sigma) gradient. The final pellet was resuspended in the same DMEM:F12 media. Cell viability was greater than 90%. Hepatocytes were plated onto primary plates (Corning). Hepatocytes were allowed to recover and attach for 4–6 h before replacement of the medium overnight before stress treatments the following day for the times and concentrations indicated.

Generation and culture of iPSC-derived human hepatocytes. The human induced pluripotent (iPS) cell line A1ATDR/R used in this work was derived as previously described^{37,38} under approval by the regional research ethics committee (reference number 08/H0311/201). iPS cells were maintained in Essential 8 chemically defined media³⁹ supplemented with 2 ng ml⁻¹ Tgf-β (R&D) and 25 ng ml⁻¹ FGF2 (R&D), and cultured on plates coated with 10 μg ml⁻¹ Vitronectin XFTM (STEMCELL Technologies). Colonies were regularly passaged by short-term incubation with 0.5 mM EDTA in PBS. For hepatocyte differentiation, colonies were dissociated into single cells following incubation with StemPro Accutase Cell Dissociation Reagent (Gibco) for 5 min at 37 °C. Single cell suspensions were seeded on plates coated with 10 μg ml⁻¹ Vitronectin XFTM (STEMCELL Technologies) in maintenance media supplemented with 10 μM ROCK Inhibitor Y-27632 (Selleckchem) and grown for up to 72 h before differentiation. Hepatocytes were differentiated as previously reported⁴⁰, with minor modifications as listed. In brief, following endoderm differentiation, anterior foregut specification was achieved after 5 days of culture with RPMI-B27 differentiation media supplemented with 50 ng ml⁻¹ Activin A (R&D)⁴⁰. Foregut cells were further differentiated into hepatocytes with HepatoZYME-SFM (Gibco) supplemented with 2 mM L-glutamine (Gibco), 1% penicillin-streptomycin (Gibco), 2% non-essential amino acids (Gibco), 2% chemically defined lipids (Gibco), 14 μg ml⁻¹ of insulin (Roche), 30 μg ml⁻¹ of transferrin (Roche),

Article

50 ng ml⁻¹ hepatocyte growth factor (R&D), and 20 ng ml⁻¹ oncostatin M (R&D), for up to 27 days.

Cellular studies on ISR

Chemicals and reagents. Tunicamycin and ISRIB were purchased from Sigma-Aldrich. Metformin and Phenformin was purchased from Cayman Chemicals and GSK2606414 from Calbiochem. The antibody for GDF15 and CHOP (sc-7351) were obtained from Santa Cruz. Phospho S51 EIF2 α (ab32157) and Calnexin (ab75801) were purchased from Abcam. D. Ron provided the antibody for ATF4.

Eukaryotic cell lines and treatments. Mouse embryonic fibroblast (MEF) cell lines were obtained from D. Ron (CIMR/IMS, Cambridge) and maintained as previously described¹⁸. MEFs were transfected with 30 nM control siRNA or a smartpool on-target plus siRNA for mouse *Chop* (Dharmacon L-062068-00-0005) using Lipofectamine RNAi MAX (Invitrogen) according to the manufacturer's instruction. 48 h post siRNA transfection, cells were processed for RNA and protein expression analysis. All cells were maintained at 37 °C in a humidified atmosphere of 5% CO₂ and seeded onto 6- or 12-well plates before stress treatments for the times and concentrations indicated. Vehicle treatments (for example, DMSO) were used for control cells when appropriate.

RNA isolation, cDNA synthesis and qPCR. Following treatments, cells were lysed with Buffer RLT (Qiagen) containing 1% 2-mercaptoethanol and processed through a Qias shredder with total RNA extracted using the RNeasy isolation kit according to manufacturer's instructions (Qiagen). RNA concentration and quality was determined by Nanodrop. 400 ng–500 ng of total RNA was treated with DNaseI (ThermoFisher Scientific) and then converted to cDNA using MMLV Reverse Transcriptase with random primers (Promega). Quantitative RT–PCR was carried out with either TaqMan Universal PCR Master Mix or SYBR Green PCR master mix on the QuantStudio 7 Flex Real time PCR system (Applied Biosystems). All reactions were carried out in either duplicate or triplicate and C_t values were obtained. Relative differences in gene expression were normalized to the expression levels of the housekeeping genes *HPRT* or *GAPDH* for cell analysis, using the standard curve method. Primers used for this study: mouse *Gdf15* (Mm00442228_m1, ThermoFisher Scientific), human *GDF15* (Hs00171132_m1, ThermoFisher Scientific), human *GAPDH* (Hs02758991_g1, ThermoFisher Scientific), mouse *Hprt* (forward AGCCTAAGATGAGCGCAAGT, reverse GGCCACAGGACTAG AACACC).

Immunoblotting. Following treatments, cells were washed twice with ice-cold D-PBS and proteins collected using RIPA buffer supplemented with cOmplete protease and PhosStop inhibitors (Sigma). The lysates were cleared by centrifugation at 13,000 rpm for 15 min at 4 °C, and protein concentration determined by a Bio-Rad DC protein assay. Typically, 20–30 μ g of protein lysates were denatured in NuPAGE 4 \times LDS sample buffer and resolved on NuPage 4–12% Bis-Tris gels (Invitrogen) and the proteins were transferred by iBlot (Invitrogen) onto nitrocellulose membranes. The membranes were blocked with 5% non-fat dry milk or 5% BSA (Sigma) for 1 h at room temperature and incubated with the antibodies described in the reagents section. Following a 16 h incubation at 4 °C, all membranes were washed five times in Tris-buffered saline/0.1% Tween-20 before incubation with horseradish peroxidase (HRP)-conjugated anti-rabbit immunoglobulin G (IgG) or HRP-conjugated anti-mouse IgG (Cell Signalling Technologies). The bands were visualized using Immobilon Western Chemiluminescent HRP Substrate (Millipore). All images were acquired on the ImageQuant LAS 4000 (GE Healthcare).

Statistical analyses

CAMERA data were analysed using a mixed linear model with restricted maximum likelihood to investigate the metformin effect on GDF15. This is analogous to conducting a repeated measures ANOVA, but is

a more flexible analysis and allows for missing observations within subjects. The 0–18 months difference in weight and GDF15 correlation was tested using Spearman's coefficient. CAMERA data were analysed using STATA v.15.1.

Other statistical analyses were performed using Prism 7 and Prism 8, using unpaired two-tailed *t*-tests, or two-way ANOVA, with multiple comparison adjustment by Tukey's or Sidak's test. Metabolic rate was determined using ANCOVA with energy expenditure as the dependent variable, body weight as a covariate and treatment as a fixed factor. ANCOVA and analyses of glucose and ITT in mice were performed using SPSS 25 (IBM).

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Source Data for Figs. 1–4 and Extended Data Figs. 1–6, 8–10 are provided with the paper. Other data that support the findings of this study are available from the corresponding authors upon request. The CAMERA trial dataset is held at the University of Glasgow and is available on request from the investigators subject to a signed agreement operating within the confines of the original ethics application.

- Skarnes, W. C. et al. A conditional knockout resource for the genome-wide study of mouse gene function. *Nature* **474**, 337–342 (2011).
- Bradley, A. et al. The mammalian gene function resource: the International Knockout Mouse Consortium. *Mamm. Genome* **23**, 580–586 (2012).
- Pettitt, S. J. et al. Agouti C57BL/6N embryonic stem cells for mouse genetic resources. *Nat. Methods* **6**, 493–495 (2009).
- McNeilly, A. D., Williamson, R., Balfour, D. J., Stewart, C. A. & Sutherland, C. A high-fat diet-induced cognitive deficit in rats that is not prevented by improving insulin sensitivity with metformin. *Diabetologia* **55**, 3061–3070 (2012).
- Ortega-Cava, C. F. et al. Strategic compartmentalization of Toll-like receptor 4 in the mouse gut. *J. Immunol.* **170**, 3977–3985 (2003).
- Goldspink, D. A. et al. Mechanistic insights into the detection of free fatty and bile acids by ileal glucagon-like peptide-1 secreting cells. *Mol. Metab.* **7**, 90–101 (2018).
- Sato, T. et al. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology* **141**, 1762–1772 (2011).
- Rashid, S. T. et al. Modeling inherited metabolic disorders of the liver using human induced pluripotent stem cells. *J. Clin. Invest.* **120**, 3127–3136 (2010).
- Yusa, K. et al. Targeted gene correction of α 1-antitrypsin deficiency in induced pluripotent stem cells. *Nature* **478**, 391–394 (2011).
- Chen, G. et al. Chemically defined conditions for human iPSC derivation and culture. *Nat. Methods* **8**, 424–429 (2011).
- Hannan, N. R., Segeritz, C. P., Touboul, T. & Vallier, L. Production of hepatocyte-like cells from human pluripotent stem cells. *Nat. Protoc.* **8**, 430–437 (2013).

Acknowledgements The CAMERA trial is funded by a project grant from the Chief Scientist Office, Scotland (CZB/4/613). D.P. is supported by a University of Oxford British Heart Foundation Centre of Research Excellence Senior Transition Fellowship (RE/13/130181). N.S. and P.W. acknowledge support from a BHF Centre of Excellence award (COE/RE/18/6/34217). We thank P. Barker, K. Burling and other members of the Cambridge Biochemical Assay Laboratory (CBAL). This project is supported by the National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. A.P.C., D. Rimmington, J.A.T., I.C., Y.C.L.T. and G.S.H.Y. are supported by the Medical Research Council (MRC Metabolic Diseases Unit (MC_UU_00014/1)). Mouse studies in Cambridge are supported by S. Grocott and the Disease Model Core, with pathology support from J. Warner and the Histopathology Core (MRC Metabolic Diseases Unit (MC_UU_00014/5) and Wellcome Trust Strategic Award (100574/Z/12/Z)). D.B.S. and S.O. are supported by the Wellcome Trust (WT 107064 and WT 095515/Z/11/Z), the MRC Metabolic Disease Unit (MC_UU_00014/1) and The National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre and NIHR Rare Disease Translational Research Collaboration. We thank J. Jones and other members of the Histopathology and ISH Core Facility, Cancer Research UK Cambridge Institute, University of Cambridge, Li Ka Shing Centre. D. Ron is supported by a Wellcome Trust Principal Research Fellowship (Wellcome 200848/Z/16/Z) and a Wellcome Trust Strategic Award to the Cambridge Institute for Medical Research (Wellcome 100140). A.V.-P., S.R.-C. and S.V. are supported by the BHF (RG/18/7/33636) and MRC (MC_UU_00014/2). A.M. is supported by a studentship from the Experimental Medicine Training Initiative/AstraZeneca. R.A.T. and L.V. are supported by an ERC advanced grant NewChol and core support from the Wellcome Trust and Medical Research Council to the Wellcome–Medical Research Council Cambridge Stem Cell Institute. M.Y., D.A.G., E.L.M., F.M.G. and F.R. are supported by the MRC (MC_UU_00014/3) and Wellcome Trust (106262/Z/14/Z and 106263/Z/14/Z). M.Y. is supported by a BBSRC-DTP studentship. A.R.K., R.R.E. and K.S.N. are supported by NIH Grants R21 AG60139, UL1 TR000135 and T32DK007352 and acknowledge K. Klaus for technical assistance. N.J.W. is supported by the MRC (MC_UU_12015/1) and is an

NIHR Senior Investigator. We acknowledge J. Luan for statistical assistance. *CHOP*-null mice were a gift from J. Goodall.

Author contributions The overall conceptualization of studies included in this work was done by A.P.C., N.S., D.B.S., B.B.A. and S.O.: these authors contributed equally to this work. A.P.C., M.C., P.T., D. Rimmington, I.C. and Y.C.L.T. designed, managed, performed and analysed data from mouse experiments. S.V. designed experiments and analysed data. A.M. and G.S.H.Y. contributed to conceptualization of experiments and data analysis. J.A.T. performed in situ hybridization experiments. S.P. designed, managed and performed cell based assays along with E.L.M., S.R.-C., R.A.T., H.P.H., A.V-P., L.V. and D. Ron. J.T.-J.H. performed measurement of serum metformin levels. M.Y., D.A.G., F.M.G. and F.R. designed, performed and analysed organoid experiments. A.R.K., R.R.E. and K.S.N. designed and performed short-term metformin studies in humans. N.J.W. analysed the Ely Study Cohort. P.W., D.P. and N.S. designed, analysed and interpreted data arising from the CAMERA study. A.P.C., D.B.S., B.B.A. and S.O. wrote the paper, which was reviewed and edited by all the authors.

Competing interests P.W. has received grant support from Roche Diagnostics, AstraZeneca and Boehringer Ingelheim. N.S. has consulted for AstraZeneca, Boehringer Ingelheim, Eli Lilly, Napp, Novo Nordisk and Sanofi, and received grant support from Boehringer Ingelheim. M.C., P.T. and B.B.A. are or were employees of NGM Biopharmaceuticals and may hold NGM stock or stock options. F.R. and F.M.G. have received support from AstraZeneca and Eli Lilly. F.M.G. has provided remunerated consultancy services to Kallyope. S.O. has provided remunerated consultancy services to Pfizer, AstraZeneca, Novo-Nordisk and ERX Pharmaceuticals. All other authors declare no competing interests.

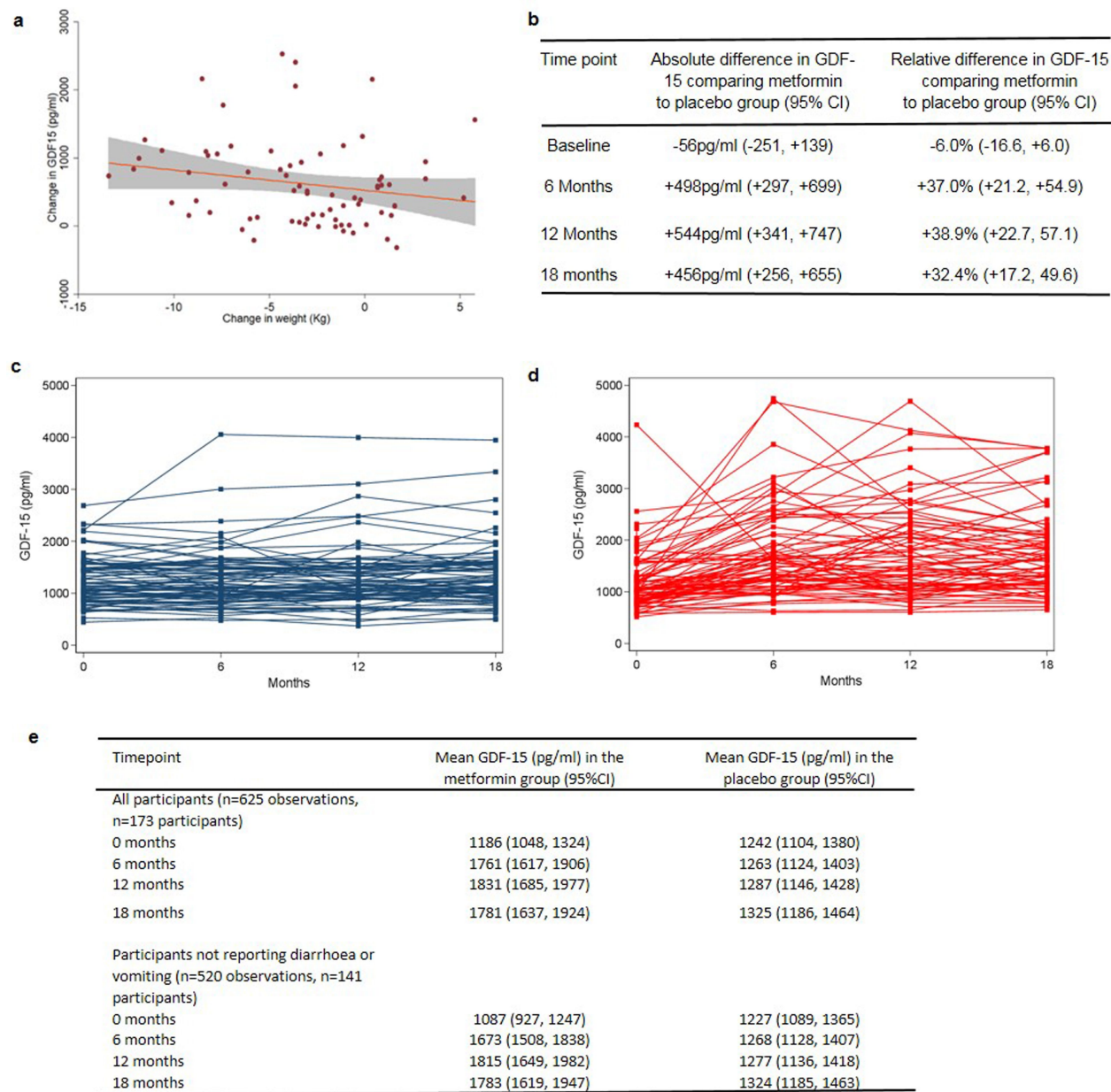
Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-019-1911-y>.

Correspondence and requests for materials should be addressed to A.P.C. or S.O.

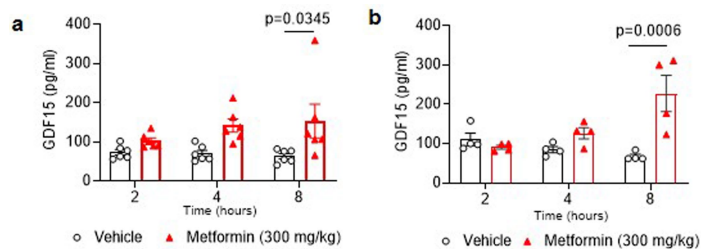
Peer review information *Nature* thanks Samuel Breit, Daniel Drucker, Jerrold M. Olefsky, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

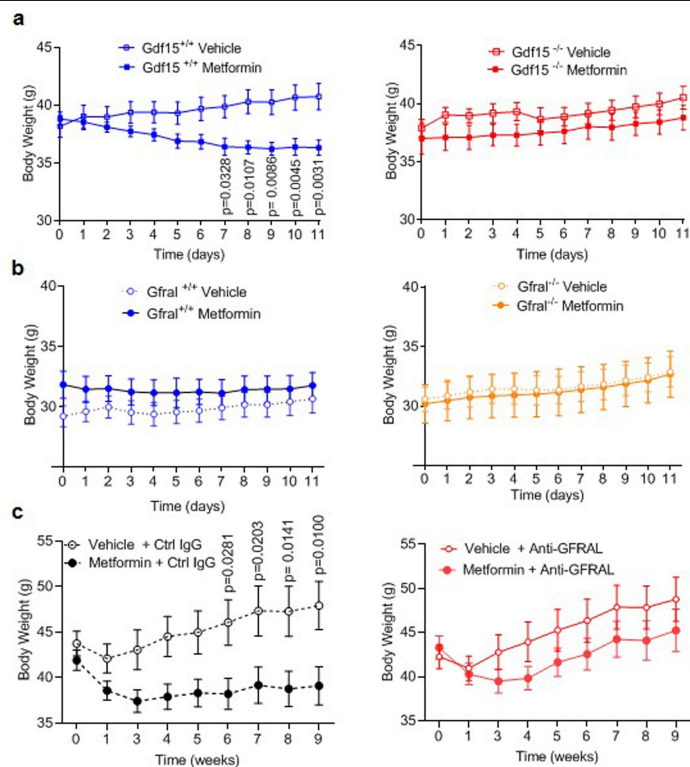


Extended Data Fig. 1 | Expanded CAMERA dataset. a, Linear association between change in body weight and change in plasma GDF15 between 0 and 18 months among metformin-treated participants ($n = 74$, Spearman correlation $r = -0.26$, two-sided $P = 0.024$). The red line is the linear regression slope, and grey area is 95% CI for the slope. **b**, Absolute and relative differences in plasma GDF15 concentration between metformin and placebo groups at each time point (total 625 observations in 173 participants). **c**, **d**, Individual measures of

plasma GDF15 levels in placebo group (**c**) and metformin group (**d**) over time. **e**, Plasma GDF15 concentration (95% CI) in overweight or obese non-diabetic participants with known cardiovascular disease randomized to metformin or placebo in CAMERA; modelled using a mixed linear model as per Fig. 1 and grouped as ‘all participants’ and ‘all participants not reporting diarrhoea and vomiting’. Model includes all participants.

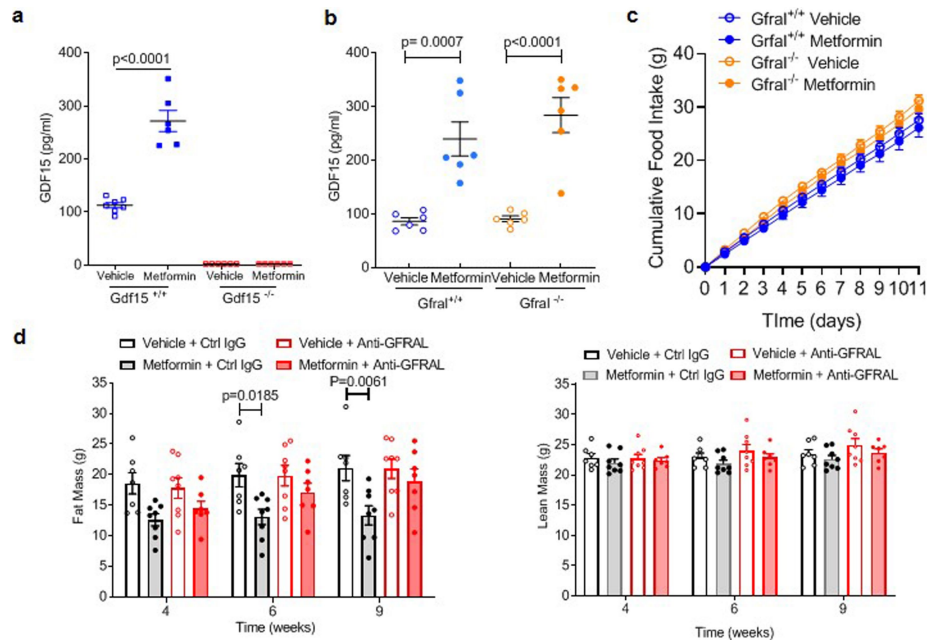


Extended Data Fig. 2 | Effect of single oral dose of metformin in chow-fed mice. Serum GDF15 levels in male mice measured 2, 4 or 8 h after a single gavage dose of metformin (300 mg kg^{-1}). **a**, Mice fed ad libitum overnight before gavage. **b**, Mice fasted for 12 h before gavage. Data are mean \pm s.e.m. (**a**; $n = 6$ per group, **b**; $n = 4$ per group); P values by two-way ANOVA with Tukey's correction for multiple comparisons.



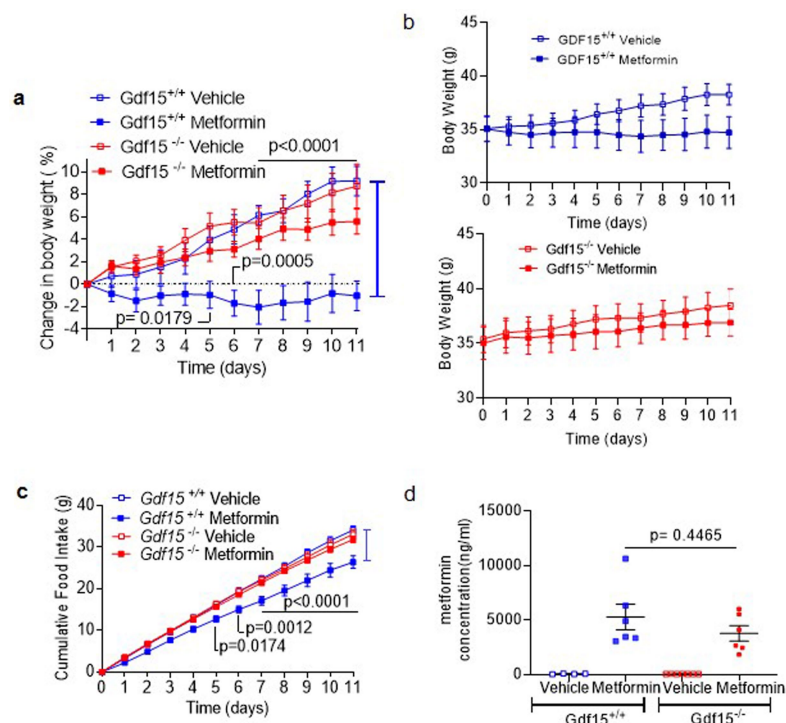
Extended Data Fig. 3 | Body weight changes with metformin treatment in mice with disrupted GDF15–GFRAL signalling. **a**, Absolute body weight in *Gdf15*^{+/+} and *Gdf15*^{-/-} mice on a high-fat diet treated with metformin (300 mg kg⁻¹ day⁻¹) for 11 days, mice as in Fig. 2a. Data are mean ± s.e.m., *P* values by two-way ANOVA with Tukey's correction for multiple comparisons. **b**, Absolute body weight in high-fat diet-fed *Gfral*^{+/+} and *Gfral*^{-/-} mice given an

oral dose of metformin (300 mg kg⁻¹) once daily for 11 days, mice as in Fig. 2c. Data are mean ± s.e.m. **c**, Absolute body weight of metformin-treated, obese mice dosed with an anti-GFRAL antagonist antibody or with control IgG weekly for five weeks, starting four weeks after initial metformin exposure; mice as in Fig. 2d. Data are mean ± s.e.m. *P* values by two-way ANOVA with Tukey's correction for multiple comparisons.



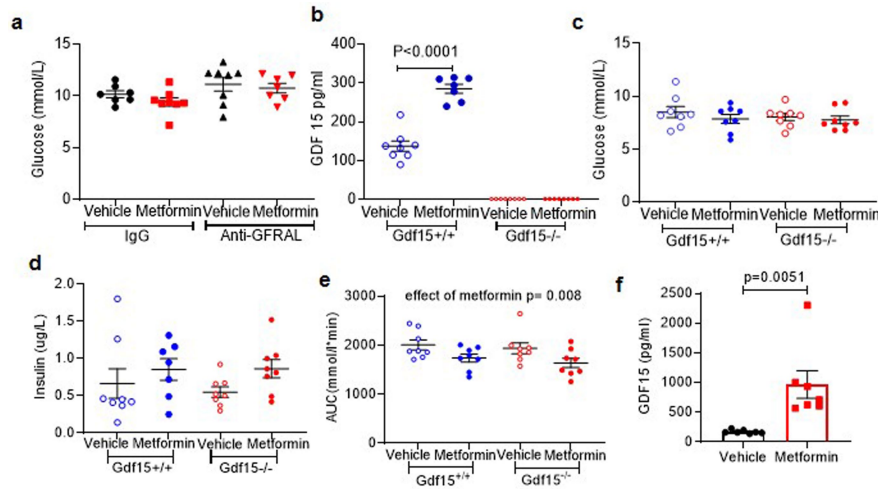
Extended Data Fig. 4 | Response of high-fat diet-fed *Gdf15*^{-/-} and *Gfral*^{-/-} mice to metformin. **a**, Circulating GDF15 levels in high-fat diet-fed *Gdf15*^{+/+} and *Gdf15*^{-/-} mice given oral dose of metformin (300 mg kg⁻¹) once daily for 11 days. Data are mean \pm s.e.m., mice as in Fig. 2a. All *Gdf15*^{-/-} samples were below lower limit of the assay (< 2 pg ml⁻¹); *P* values by two-way ANOVA with Tukey's correction for multiple comparisons. **b**, Circulating GDF15 levels in high-fat diet-fed *Gfral*^{+/+} and *Gfral*^{-/-} mice given oral dose of metformin (300 mg kg⁻¹) once daily for 11 days. Data are mean \pm s.e.m., mice as in Fig. 2c; *P* values by two-way ANOVA with Tukey's correction for multiple comparisons. **c**, Cumulative food intake in high-fat diet fed *Gfral*^{+/+} and *Gfral*^{-/-} mice on a high-fat diet given an oral dose of metformin (300 mg kg⁻¹) once daily for 11 days. Data are

mean \pm s.e.m., mice as in Fig. 2c; no statistically significant difference in vehicle versus metformin by two-way ANOVA. **d**, Fat mass (left) and lean mass (right) in metformin-treated obese mice dosed with anti-GFRAL antagonist antibody weekly for five weeks, starting four weeks after initial metformin exposure (mice as in Fig. 2d). Body composition was measured using MRI after 4 weeks of metformin exposure, before receiving anti-GFRAL (week 4), after 6 weeks of metformin exposure and 2 weeks after receiving anti-GFRAL (week 6) and after 9 weeks of metformin exposure and 5 weeks after receiving anti-GFRAL (week 9). Data are mean \pm s.e.m. ($n = 7$, vehicle + control IgG and metformin + anti-GFRAL; $n = 8$, other groups); *P* values by two-way ANOVA with Tukey's correction for multiple comparisons.



Extended Data Fig. 5 | Response of second, independent cohort of high-fat diet fed $Gdf15^{+/+}$ and $Gdf15^{-/-}$ mice to metformin. a–c. Percentage change in body weight (a), absolute body weight (b) and cumulative food intake (c) of $Gdf15^{+/+}$ and $Gdf15^{-/-}$ mice on a high-fat diet treated with metformin ($300 \text{ mg kg}^{-1} \text{ day}^{-1}$) for 11 days. Data are mean \pm s.e.m. ($n = 6$ per group, except

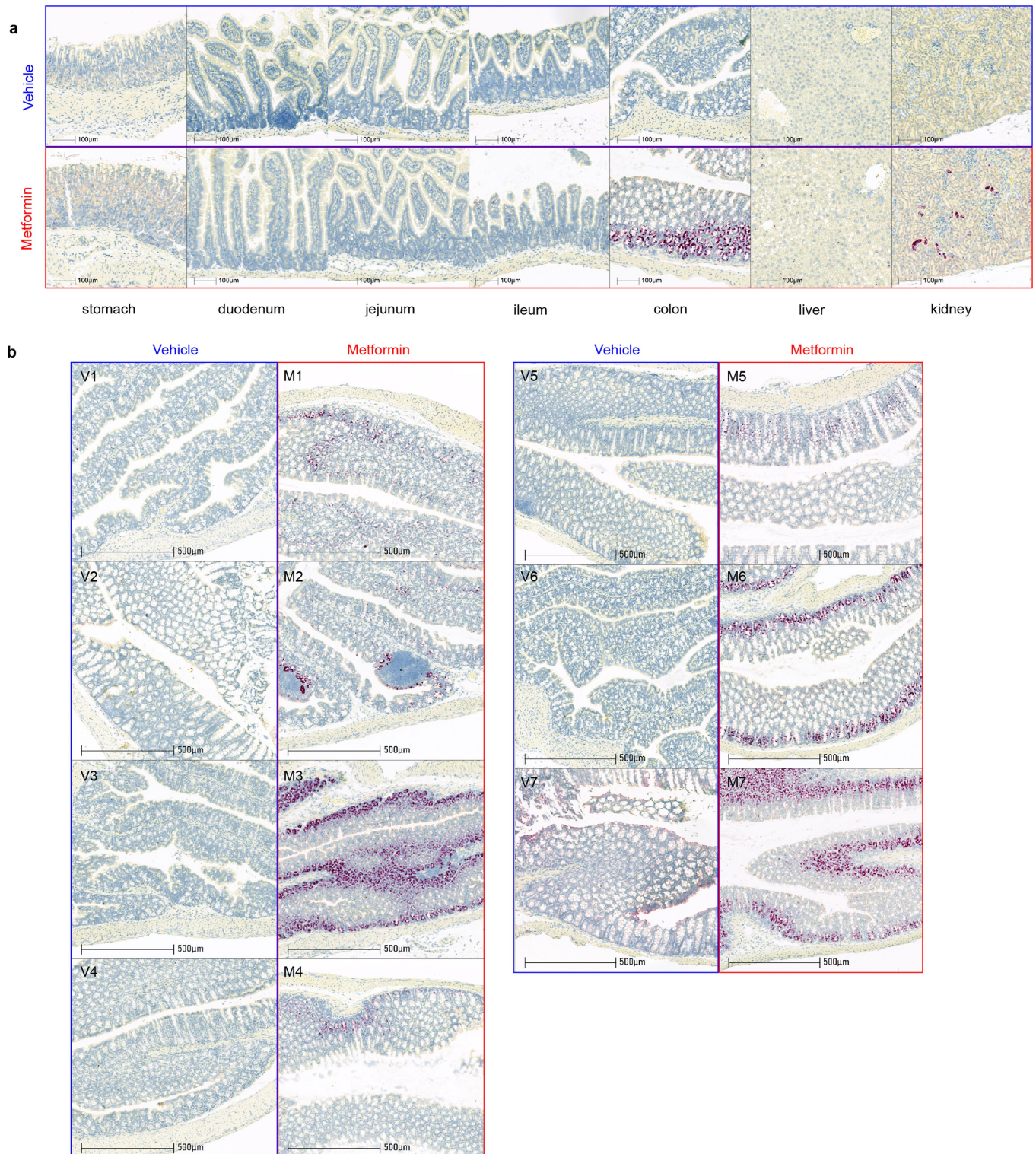
$Gdf15^{-/-}$ vehicle, $n = 7$); P values by two-way ANOVA with Tukey's correction for multiple comparisons. **d.** Circulating metformin levels in mice 6 h after final dose of metformin on day 11. Data are mean \pm s.e.m. ($n = 6$ per group, except $Gdf15^{+/+}$ vehicle, $n = 4$; $Gdf15^{-/-}$ vehicle, $n = 7$); P values by two-way ANOVA with Tukey's correction for multiple comparisons.



Extended Data Fig. 6 | Glucose, insulin and GDF15 response to metformin.

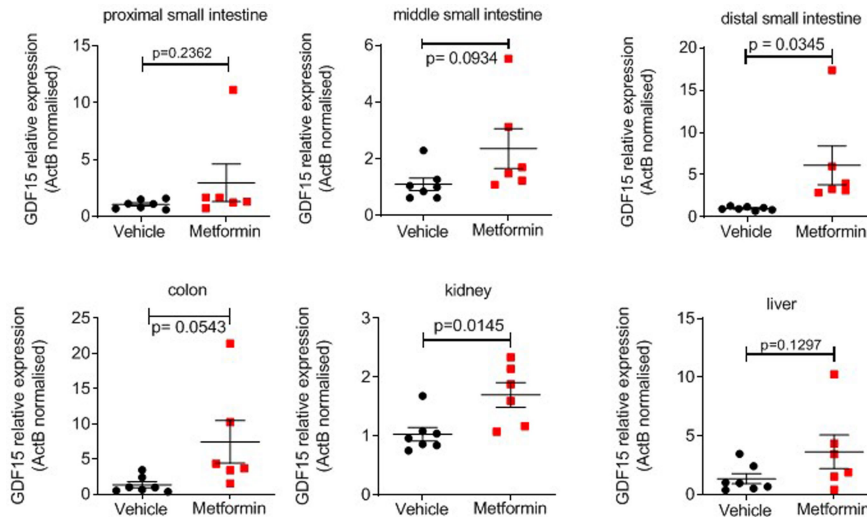
a, Fasting glucose from oral GTT as in Fig. 3e, f. ANOVA; effect of antibody, $P = 0.028$; effect of metformin, $P = 0.271$; interaction of antibody and metformin, $P = 0.707$. **b**, Circulating GDF15 in mice undergoing intraperitoneal GTT after a single dose of metformin as in Fig. 3k, l. P values by two-way ANOVA with Tukey's correction for multiple comparisons. **c**, **d**, Fasting glucose (**c**) and fasting insulin (**d**) at time 0 of intraperitoneal GTT as in Fig. 3k, l; not

statistically significant by two-way ANOVA. **e**, AUC analysis of glucose levels as in Fig. 3k, l. P values by two-way ANOVA, effect of genotype, $P = 0.392$; interaction of genotype and metformin, $P = 0.883$. **a–e**, Data all mean \pm s.e.m. **f**, Circulating GDF15 levels in high-fat-diet-fed *Gdf15*^{+/+} mice after single oral dose of metformin (600 mg kg⁻¹). Samples were collected 6 h after dosing, data are mean \pm s.e.m., $n = 7$ per group; P values (95% CI) by two-tailed t -test.

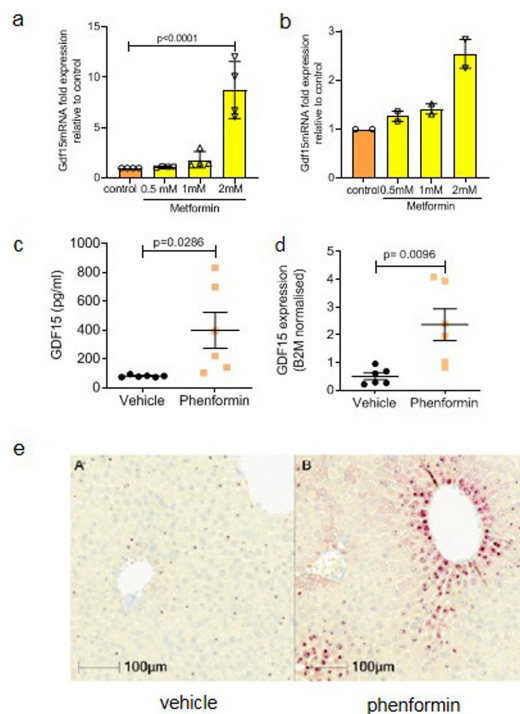


Extended Data Fig. 7 | In situ hybridization for *Gdf15* mRNA expression in gut, liver and kidney. a, Representative images from the mouse with circulating GDF15 level closest to the group median shown in Fig. 4b, with images from other regions of the gut and from liver. **b,** In situ hybridization for

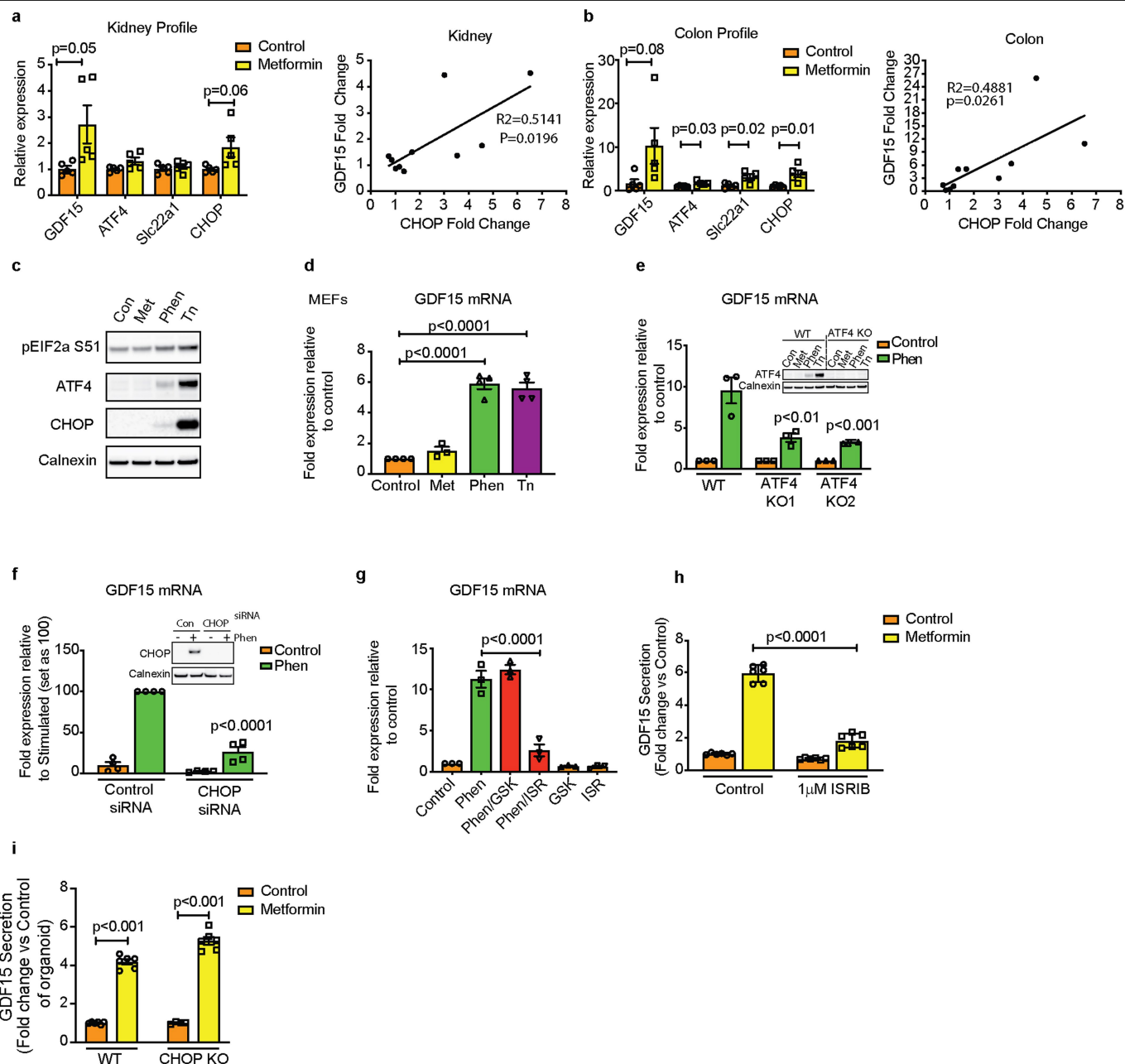
Gdf15 mRNA expression (red spots) in colon. Tissue collected from high-fat-diet-fed wild-type mice, 6 h after a single dose of oral metformin (600 mg kg^{-1}) (right, red box, M1–M7) or vehicle gavage (left, blue box, V1–V7); $n = 7$ mice per group, mice as in Fig. 4.



Extended Data Fig. 8 | Analysis of *Gdf15* mRNA expression (normalized to expression levels of *Actb*) in tissue from high-fat diet-fed *Gdf15*^{+/+} mice. Metformin treatment (300 mg kg⁻¹) once daily for 11 days (see Fig. 2a). Data are mean ± s.e.m., *n* = 6 metformin, *n* = 7 vehicle; *P* values (95% CI) by two-tailed *t*-test.



Extended Data Fig. 9 | Hepatic GDF15 response to biguanides. a, b, *Gdf15* mRNA expression in primary mouse hepatocytes (a) or human iPS-cell-derived hepatocytes (b) treated with vehicle control (Con) or metformin for 6 h. mRNA expression is presented as fold expression relative to control treatment (set at 1), normalized to *Hprt* and *GAPDH* in mouse and human cells, respectively. Data are expressed as mean \pm s.e.m. from four (a) or two (b) independent experiments. *P* values (95% CI) by one-way ANOVA with Tukey's correction for multiple comparisons. c, d, Circulating levels of GDF15 (c) and hepatic *Gdf15* mRNA expression (d) (normalized to β 2-microglobulin) in chow-fed, wild-type mice 4 h after a single oral dose of phenformin (300 mg kg^{-1}). Data are mean \pm s.e.m., $n = 6$ per group; *P* values (95% CI) by two tailed *t*-test. e, Representative image of in situ hybridization for *Gdf15* mRNA expression (red spots) of fixed liver tissue derived from animals treated as described in c and d.



Extended Data Fig. 10 | Role of the ISR in biguanide-induced *Gdf15* expression.

a, b, mRNA levels in kidney (**a**) and colon (**b**) isolated from obese mice 24 h after a single oral dose of metformin (600 mg kg⁻¹). Data are mean \pm s.e.m. ($n=5$ per group, except for $n=4$ for colon metformin Slc22a1). *P* values (95% CI) by two-tailed *t*-test. *Gdf15* mRNA fold induction 24 h after metformin (600 mg kg⁻¹) is positively correlated with *Chop* mRNA induction in both kidney (**a**, right) and colon (**b**, right). Black line shows linear regression analysis. **c–g**, Immunoblot analysis of ISR components (**c**) and *Gdf15* mRNA expression (**d**) in wild-type mouse embryonic fibroblasts (MEF) treated with vehicle control (Con), metformin (Met, 2 mM) or phenformin (Phen, 5 mM) or tunicamycin (Tn, 5 μ g ml⁻¹, used as a positive control) for 6 h. **e–g**, *Gdf15* mRNA expression in ATF4 knockout (KO) MEFs (**e**), in control siRNA and *CHOP* siRNA transfected wild-type MEFs treated with Tn or Phen for 6 h (**f**), or in wild-type MEFs pre-treated for 1 h with either the PERK inhibitor GSK2606414 (GSK, 200 nM) or eIF2 α inhibitor ISRIB (ISR, 100 nM) then co-treated with

phenformin for a further 6 h (**g**). mRNA expression is presented as fold expression relative to its respective control treatment (set as 1) or phenformin-treated samples (set as 100) with normalization to *Hprt* gene expression. Data are mean \pm s.e.m. from two (**c, d**) or at least three (**e–g**) independent experiments. *P* values (95% CI) by two-tailed *t*-test relative to phenformin-treated control wild-type and control siRNA-treated samples. **h**, GDF15 protein in supernatant of mouse derived 2D duodenal organoids treated with metformin in the absence or presence of ISRIB (1 μ M). Data are expressed as mean \pm s.e.m. from two independent experiments. At least duplicate protein measurements for each sample. *P* values by two-way ANOVA with Sidak's correction for multiple comparisons. **i**, GDF15 protein in supernatants of mouse-derived 2D duodenal organoids from wild-type and *Chop*-null mice treated with metformin from two independent experiments. At least duplicate protein measurements for each sample. Data are mean \pm s.e.m.; *P* values (95% CI) by two-tailed *t*-test.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | | |
|-------------------------------------|--|
| n/a | Confirmed |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of all covariates tested |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	No software was used.
Data analysis	Prism v8 for MacOSX , Prism V 8 for Windows and Microsoft Excel 13 were used for Statistical analysis. CAMERA data analysed using STATA version 15.1. ANCOVA and analysis of data from glucose homeostasis experiments was performed using SPSS 25 (IBM).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are available from the corresponding authors upon request. The CAMERA trial dataset is held at the University of Glasgow and is available on request from the investigators subject to a signed agreement operating within the confines of the original ethics application.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were determined on the basis of homogeneity and consistency of characteristics in the selected models and were sufficient to detect statistically significant differences in body weight, food intake and serum parameters between groups, while also ensuring no more animals than necessary were used. These assessments are based on extensive expertise with these models and endpoints. In vitro sample sizes were based on previous extensive experience with reagents and systems used.
Data exclusions	One data point of 25 food intake points collected on day 11 of mouse study 3 was lost due to technical error (see methods). One animal in mouse study 6 which was otherwise well and healthy did not progress to ITT after metformin due to technical and acute behavioural issues at time of study. One blood sample from mouse study 7 was lost due to technical error.
Replication	Acute dosing of metformin to chow fed animals has been replicated and reproduced across two laboratories. Phenformin data has been replicated. Mouse study 1 has not been replicated exactly but the acute response to metformin in HFD fed animals has been reproduced across two laboratories. Longer term dosing studies of metformin to HFD Gf15 and Gdf15 null mice (mouse study 3 and 4) have been replicated; two independent cohorts of Gdf15 null mice appear in the manuscript. ITT post metformin dosing of Gdf15 null has been done once. Mouse study 5 has not been replicated exactly but similar data were generated using anti-GFRAL antibody given over a different time scale but with similar metformin exposure. Mouse studies 7 and 8 have been done once. All in vitro and cell based experiments have been replicated successfully and reliably reproduced with replicate numbers reported in legends.
Randomization	Animals were randomised into treatment groups based on body weight such that the mean body weights of each group were as matched as possible but without using excess numbers of animals.
Blinding	Serum/plasma GDF15 measurements were blinded by the investigators. Investigator undertaking ISH analysis of tissue was blinded to treatment during tissue processing and labelling.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input type="checkbox"/>	<input checked="" type="checkbox"/> Clinical data

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used	ATF4 antibody was obtained from David Ron. CHOP was obtained from Santa Cruz (Cat# sc-7351; RRID: AB_627411). pEIF2a Ser51 was obtained from Abcam (Cat# ab32157; RRID: AB_732117). Calnexin was obtained from Abcam (Cat# ab75801; RRID: AB_1310022). Anti-GFRAL functional blocking antibody generated by NGM. Secondary antibodies used were horseradish peroxidase (HRP)-conjugated anti-rabbit immunoglobulin G (IgG), HRP-conjugated anti-mouse IgG (Cell Signalling Technologies)
Validation	ATF4 and CHOP antibody has been previously validated in KO and siRNA knockdown samples (this study and PMID: 30639358). pEIF2a antibody has been validated previously (PMID: 27297692) and we have independent confirmed this during the course of our studies for PMID: 30639358 (data not shown). Calnexin is a well established commercially available antibody that has been used by numerous investigators and published frequently. Characterization of anti-GFRAL functional blocking antibodies was carried out using a cell-based RET/GFRAL luciferase gene reporter assays, in vitro binding studies (ELISA and Biacore) and in vivo studies as described in patent number; US10174119B2, https://patents.google.com/patent/US10174119B2/en .

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	Mouse Embryonic Fibroblasts (MEFs) was obtained David Ron (CIMR). The hiPSC line A1ATDR/R was obtained from Ludovic Vallier (Cambridge Stem Cell Institute).
Authentication	MEFs have been previously validated (PMID 12667446 and 12667446) and the hiPS cell line validated (PMID 20739751 and 21993621)
Mycoplasma contamination	MEFs and A1ATDR/R cells were tested negative for Mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	no commonly misidentified lines were used.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	<p>Mice studied ranged between 2 to 6 months of age. At NGM, animals were kept under controlled light (12hour light and 12hour dark cycle, dark 6:30 pm - 6:30 am), temperature ($22 \pm 3^{\circ}\text{C}$) and humidity ($50\% \pm 20\%$) conditions. In Cambridge, were maintained in a 12-hour light/12-hour dark cycle (lights on 0700–1900), temperature-controlled (22°C) facility, with ad libitum access to food (RM3(E) Expanded chow, Special Diets Services, UK) and water. Any mice bought from an outside supplier were acclimatised in a holding room for at least one week prior to study. Study diets as outlined in relevant section of "Methods". Age, sex and body weight of mice used as detailed in relevant "Methods" section. Unless otherwise stated, male mice were used a danimals housed singly during studies with environmental enrichment within cages.</p> <p>All drug administration and testing performed during the light cycle.</p> <p>Wild type mice used in Cambridge (C57BL6/J mice) from Charles River, Margate, UK.</p> <p>C57BL/6N-Gdf15tm1a(KOMP)Wtsi/H mice were obtained from the MRC Harwell Institute, UK.</p> <p>Gfrr1^{-/-} mice were purchased from Taconic (#TF3754) on a mixed 129/SvEv-C57BL/6 background and backcrossed for 10 generations to >99% C57BL/6 background at NGM's animal facility.</p>
Wild animals	The study did not involve wild animals
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	<p>At NGM, all experiments were conducted with NGM IACUC approved protocols and all relevant ethical regulations were complied with throughout the course of the studies, including efforts to reduce the number of animals used.</p> <p>In Cambridge, all mouse studies were performed in accordance with UK Home Office Legislation regulated under the Animals (Scientific Procedures) Act 1986 Amendment, Regulations 2012, following ethical review by the University of Cambridge Animal Welfare and Ethical Review Body (AWERB).</p>

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	<p>CAMERA was a randomized, double-blinded, placebo-controlled trial designed to investigate the effect of metformin on surrogate markers of cardiovascular disease in patients without diabetes, aged 35 to 75, with established coronary heart disease and a large waist circumference ($\geq 94\text{cm}$ in men, $\geq 80\text{cm}$ in women). Mean age (yrs)\pmSD; metformin 63(8), placebo 64(8); male sex, metformin 79(81%), placebo 63(72%). Baseline characteristics did not differ substantially between treatment groups. Nine participants completed the study by Konopka and colleagues; 7 had a family history of T2DM and 8 were metformin naïve. One participant had previously used metformin but discontinued more than 2 years before the study commenced.</p>
Recruitment	<p>3000 potential participants were identified from electronic searches of Glasgow general practice databases, supplemented by patients from hospital cardiology clinics. Of those invited, 805 replied and 356 were screened. 173 were enrolled and randomly assigned (86 to metformin, 87 to placebo). Key eligibility criteria included the use of statin therapy, history of coronary heart disease, large waist circumference ($\geq 94\text{cm}$ in men, $\geq 80\text{cm}$ in women), and no history or biochemical evidence of type 2 diabetes. Eligible participants were randomly assigned to metformin or placebo (1:1) with a randomisation sequence generated independently by computer with permuted blocks of four without stratification.</p> <p>In the study by Konopka and colleagues, inclusion criteria were: obesity (body mass index $>30\text{ kg/m}^2$), sedentary (<1 hour of structured activity per week), nonsmoking, and not taking any medication to control blood glucose.</p>
Ethics oversight	<p>All participants in CAMERA study provided written informed consent and were followed up for 18 months. This study was approved by the Medicines and Healthcare Products Regulatory Agency and West Glasgow Research Ethics Committee, and done in accordance with the principles of the Declaration of Helsinki and good clinical practice guidelines. The study by Konopka and colleagues was approved by the Mayo Clinic Institutional Review Board and all participants provided written, informed consent.</p>

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration	The CAMERA trial is registered with ClinicalTrials.gov, number NCT00723307. The study by Konopka and colleagues is registered, number NCT01956929.
Study protocol	Trial protocol outlined in PMID: 24622715 and supplied on submission.
Data collection	<p>Participants were randomized 1:1 to 850mg metformin or matched placebo twice daily with meals. Follow up study visits were conducted between 2009 and 2012 at the Glasgow Clinical Research Centre. Participants attended six monthly visits after overnight fasts and before taking their morning dose of metformin. Blood samples collected during the trial were centrifuged at 4 degrees Celsius soon after sampling, separated and stored at -80°C. Bodyweight, body fat (by bio-impedance with a Tanita BIA body fat analyser [Tanita Corporation, Tokyo, Japan]), waist circumference (measured midway between lowest rib and iliac crest), and hip circumference (measured around widest part of the buttocks) were measured at each visit.</p> <p>In the study by Konopka and colleagues, placebo or metformin (week 1, 500mg twice daily; week, 2 1000mg twice daily) were administered following a six week period of washout. Samples were collected in the morning after overnight fasting.</p>
Outcomes	<p>The primary endpoint was progression of mean distal carotid intima-media thickness (cIMT) over 18 months in the modified intention-to-treat population. Further descriptions and results for the pre-specified outcomes for the CAMERA trial are provided in a previous publication PMID: 24622715.</p> <p>As outlined in a previous publication (PMID: 27160898) the study by Konopka and colleagues aimed to investigate whether metformin inhibited glucagon-stimulated endogenous glucose production (EGP) in humans. The study measured EGP using stable isotope methodology under basal, glucagon-deficient, and glucagon-stimulated conditions.</p>

Loss of p53 drives neuron reprogramming in head and neck cancer

<https://doi.org/10.1038/s41586-020-1996-3>

Received: 3 December 2018

Accepted: 15 January 2020

Published online: 12 February 2020

 Check for updates

Moran Amit^{1,15}✉, Hideaki Takahashi^{1,2,15}, Mihnea Paul Dragomir³, Antje Lindemann¹, Frederico O. Gleber-Netto¹, Curtis R. Pickering¹, Simone Anfossi³, Abdullah A. Osman¹, Yu Cai¹, Rong Wang¹, Erik Knutsen^{3,4}, Masayoshi Shimizu^{3,5}, Cristina Ivan^{3,5}, Xiayu Rao⁶, Jing Wang⁶, Deborah A. Silverman⁷, Samantha Tam¹, Mei Zhao¹, Carlos Caulin^{8,9}, Assaf Zinger^{10,11}, Ennio Tasciotti^{10,11}, Patrick M. Dougherty¹², Adel El-Naggar¹³, George A. Calin^{3,5}✉ & Jeffrey N. Myers^{1,14}✉

The solid tumour microenvironment includes nerve fibres that arise from the peripheral nervous system^{1,2}. Recent work indicates that newly formed adrenergic nerve fibres promote tumour growth, but the origin of these nerves and the mechanism of their inception are unknown^{1,3}. Here, by comparing the transcriptomes of cancer-associated trigeminal sensory neurons with those of endogenous neurons in mouse models of oral cancer, we identified an adrenergic differentiation signature. We show that loss of *TP53* leads to adrenergic transdifferentiation of tumour-associated sensory nerves through loss of the microRNA miR-34a. Tumour growth was inhibited by sensory denervation or pharmacological blockade of adrenergic receptors, but not by chemical sympathectomy of pre-existing adrenergic nerves. A retrospective analysis of samples from oral cancer revealed that p53 status was associated with nerve density, which was in turn associated with poor clinical outcomes. This crosstalk between cancer cells and neurons represents mechanism by which tumour-associated neurons are reprogrammed towards an adrenergic phenotype that can stimulate tumour progression, and is a potential target for anticancer therapy.

Early in cancer development, nerve fibres form and infiltrate tumour tissue, and the density of these nerves in solid tumours has been associated with poor clinical outcomes^{1,2}. Neurogenic signals, including adrenergic stimulation by tumour-infiltrating neural fibres, are strongly associated with tumorigenesis, angiogenesis, invasion, and metastasis^{1,3}. Hence, the little-understood molecular mechanisms of cancer–nerve crosstalk during cancer-associated neural infiltration represents opportunities for therapeutic intervention. The *TP53* tumour suppressor gene is the most commonly mutated gene in head and neck cancer and shapes multiple aspects of tumour formation, including the microenvironment^{4,5}. p53 expression fluctuates during nerve regeneration, permitting tight control of the plasticity of the differentiation phase^{6,7}. We therefore explored the mechanisms of cancer–nerve interactions in head and neck squamous cell carcinoma on the basis of the hypothesis that the p53 protein suppresses cancer–nerve interactions in this disease and that the loss

of this p53 function increases cancer–nerve crosstalk and thereby promotes tumour progression.

Loss of p53 alters neural milieu

To evaluate the impact of tumour innervation in head and neck cancer, we analysed survival data from The Cancer Genome Atlas. High neural density in oral cavity squamous cell carcinoma (OSCC) was associated with poorer overall survival ($P < 0.0001$, log-rank test) and with the presence of *TP53* mutations ($P < 0.0001$, Extended Data Fig. 1a–c). To study the interplay between nerves and epithelial cell p53 function throughout tumorigenesis, we evaluated neuritogenesis over the course of the progression from precursor lesions to high-grade lesions in *Trp53*^{flox/flox} wild-type control and *Krt5*^{Cre}*Trp53*^{flox/flox} (*Trp53*^{−/−} in the recombined epithelial tissue) mouse models of oral cancer⁸ (Fig. 1a). Nerve density was significantly greater in tumours from mice lacking

¹Department of Head and Neck Surgery, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ²Department of Otorhinolaryngology Head and Neck Surgery, Yokohama City University, Yokohama, Japan. ³Department of Experimental Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ⁴Department of Medical Biology, Faculty of Health Sciences, UiT, The Arctic University of Norway, Tromsø, Norway. ⁵Center for RNA Interference and Non-coding RNAs, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ⁶Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ⁷Department of Melanoma Medical Oncology, Division of Cancer Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ⁸Department of Otolaryngology, Head and Neck Surgery, University of Arizona, Tucson, AZ, USA. ⁹University of Arizona Cancer Center, University of Arizona, Tucson, AZ, USA. ¹⁰Regenerative Medicine Program, Houston Methodist Research Institute, Houston, TX, USA. ¹¹Department of Orthopedics and Sports Medicine, Houston Methodist Hospital, Houston, TX, USA. ¹²Department of Pain Medicine, Division of Anesthesiology, Critical Care, and Pain Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ¹³Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ¹⁴Department of Cancer Biology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ¹⁵These authors contributed equally: Moran Amit, Hideaki Takahashi. ✉e-mail: mamit@mdanderson.org; gcalin@mdanderson.org; jmyers@mdanderson.org

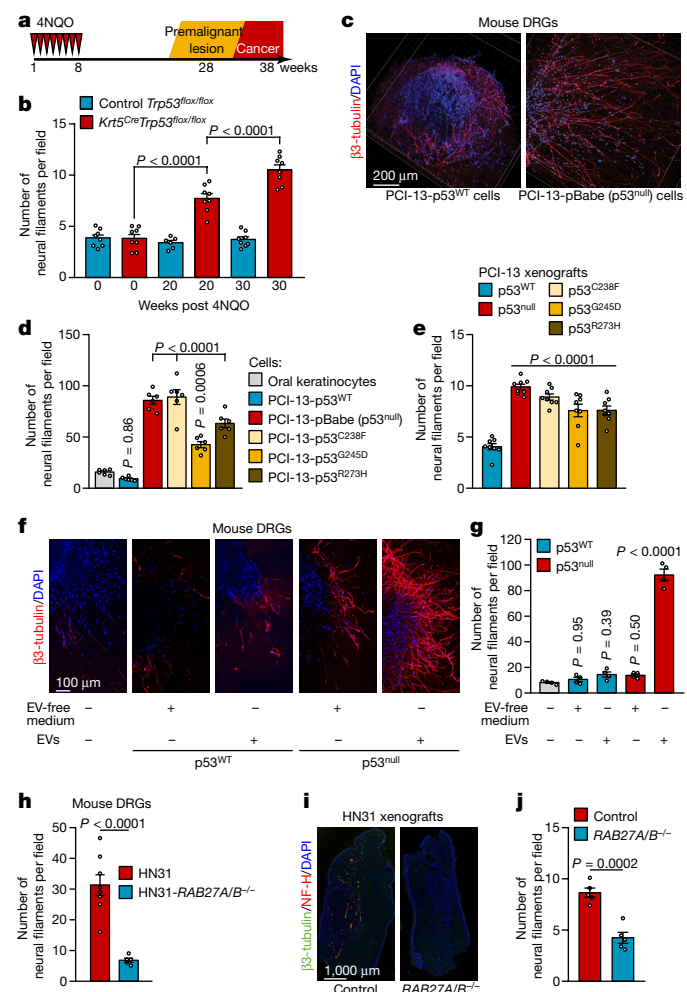


Fig. 1 | Loss of p53 alters the neural microenvironment throughout tumour evolution. **a**, Tumour progression in *Krt5.Cre;B6.129P2-Trp53^{tm1Bnn/J}* (*Krt5^{Cre}Trp53^{flax/flax}*) and *B6.129P2-Trp53^{tm1Bnn/J}* (*Trp53^{flax/flax}*) control mice. **b**, Quantification of neural density in tongues from *Trp53^{flax/flax}* and *Krt5^{Cre}Trp53^{flax/flax}* mice immediately after the end of treatment with the carcinogen 4-nitroquinoline 1-oxide (4NQO) (normal mucosa), 20 weeks after treatment completion (low-grade lesions), and 30 weeks after treatment completion (high-grade lesions) ($n = 8$ except for control group at 20 weeks ($n = 6$)). **c**, Representative immunofluorescence staining of DRGs co-cultured with p53^{WT} or p53^{null} PCI-13 cells; data independently replicated in 12 ganglia. **d**, Quantification of neurite density in DRGs co-cultured with p53-isogenic PCI-13 cells or normal oral keratinocytes ($n = 6$ biologically independent ganglia per cell line). **e**, Analysis of neural density in orthotopic p53-isogenic PCI-13 xenograft cells ($n = 8$ mice per group). **f**, Representative immunofluorescence of in vitro neurite density in DRGs treated with soluble factors (EV-depleted conditioned medium) or EVs from p53-isogenic PCI-13 cells; data independently replicated in 20 ganglia. **g**, In vitro quantification of neurite density ($n = 4$ biologically independent ganglia per condition). **h**, Quantification of neurite density in freshly collected DRG cultured with conditioned medium from HN31 RAB27A/RAB27B-isogenic human OSCC cells ($n = 8$ and $n = 5$ biologically independent ganglia for HN31 RAB27A^{+/+}RAB27B^{+/+} and HN31 RAB27A^{+/+}RAB27B^{-/-} cells respectively). **i**, Representative immunofluorescence montage of glossectomy specimens derived from mice 3 weeks after injection of HN31 RAB27A/RAB27B-isogenic OSCC cells; data independently replicated in 10 mice. **j**, In vivo analyses of neural density ($n = 5$ mice per group). Unpaired two-tailed *t*-test (**h, j**) and one-way ANOVA with Tukey multiple comparisons (**b, d, e, g**). Bar graphs represent mean \pm s.e.m.

p53 expression in oral epithelia compared with control mice expressing p53 (Fig. 1b), in both early and advanced phases of cancer development. Similarly, in orthotopic xenografts of human OSCC cells, we found

increased nerve density p53-deficient (p53^{null}) tumours compared with wild-type (p53^{WT}) controls (Extended Data Fig. 1d–f). These findings suggest that the loss of p53 signalling in epithelial tumour cells is associated with neuritogenesis during early tumorigenesis.

To better understand the mechanisms that link epithelium-derived signalling, altered p53 function, and neuritogenesis, we assessed induced neurite outgrowth from dorsal root ganglia (DRG) co-cultured with oral keratinocytes and p53-isogenic OSCC cells. Neurite outgrowth was markedly enhanced in the presence of p53^{null} cells compared with p53^{WT}-expressing cells (Fig. 1c, d). The p53^{R273H} mutation (which causes a functional p53 deficiency) was also associated with increased neuritogenesis, as was the p53^{C238F} mutation (which leads to structural p53 deficiency); however, the partial effect of p53^{G245D} on the DNA-binding region (which causes partial functional p53 deficiency) only modestly increased neuritogenesis⁹. The increased neuritogenesis seen in tumours with p53^{null} or either DNA-contact or structural mutations in *TP53*, both in vitro and in vivo (Fig. 1e, Extended Data Fig. 1g–i), suggested that the effect of p53 mutations on neuritogenesis is likely to result from loss of p53 function.

Cancer-derived vesicles control neuritogenesis

To investigate the route by which neurotrophic cues are delivered to nerves, we incubated DRGs with cancer-derived soluble and extracellular vesicle (EV) compartments of conditioned medium from p53-isogenic OSCC cell cultures (Extended Data Fig. 2a–g). DRGs cultured with EVs derived from p53^{null} cells had more neurofilaments than those cultured with the corresponding EV-depleted conditioned medium or with EVs from p53^{WT} cells (Fig. 1f, g).

To assess the potential effect of EVs on neuritogenesis, we used conditioned medium derived from HN31 cells (a human OSCC cell line with endogenous p53^{C176F} and p53^{A161S} mutations that cause functional deficiency) and isogenic HN31 RAB27A^{-/-}RAB27B^{-/-} cells lacking the GTPases RAB27A and RAB27B, which are necessary for the exocytosis of EVs¹⁰ (Extended Data Fig. 2h, i). Knockout of RAB27A and RAB27B significantly reduced DRG neuritogenesis in vitro (Fig. 1h, Extended Data Fig. 2j, k) and in vivo in orthotopic xenografts of OSCC cells (Fig. 1i, j). Thus, neurons appear to be a major stromal target of EVs derived from p53-deficient cancer cells.

Cancer-derived miRNAs drive axonogenic switch

Small RNAs are key regulators of neuronal development, regeneration, and function^{11,12}, and EVs are the main route by which RNA species are transferred between cells (Extended Data Fig. 3a, b). By comparing the microRNA (miRNA) profiles of EVs derived from p53-sufficient and isogenic p53-deficient cell lines, we identified p53-related alterations in the expression of 17 miRNAs (Extended Data Fig. 3c–e). A customized single-channel Agilent array revealed a similar significant increase in the expression of miR-34a-5p and miR-141-5p, but not the other 15 miRNAs, in EVs derived from p53^{WT} cells compared with those from p53^{null} cells (GEO accession number GSE140324). Analysis of miR-34a and miR-141 in tongue tumour specimens from *Trp53^{flax/flax}* control mice and *Krt5^{Cre}Trp53^{flax/flax}* mice confirmed a significant decrease in the levels of both miRNAs in *Trp53^{-/-}* tumours (Extended Data Fig. 3f).

To investigate how miR-34a and miR-141 interact with other cancer-derived EV signals to mediate axonogenesis, we co-cultured trigeminal ganglion (TG) neurons transfected with antagomiR-34a or antagomiR-141 (Extended Data Fig. 3g) with EVs containing abundant miR-34a and miR-141 derived from p53^{WT} OSCC cells. Compared with non-specific antagomiR or no-EV controls, both antagomiRs increased the number of neurofilaments, although only antagomiR-34a had a significant effect (Fig. 2a, b). Conversely, transfection of TG neurons with ectopic miR-34a (compared with scramble miRNA or no-EV controls), followed by co-culture with miR-34a-deficient EVs derived

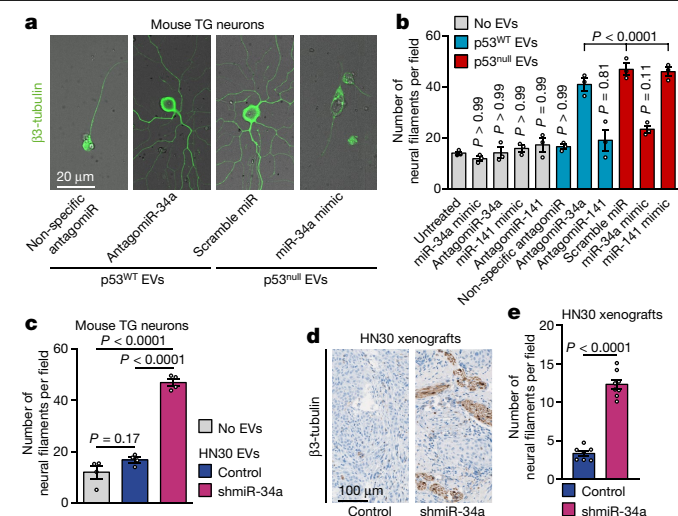


Fig. 2 | p53-dependent alterations in miRNA populations control neuritogenesis. **a**, Representative fluorescence–bright-field overlay images of TG neurons after transfection with antagomiR-34a and incubation with EVs from p53^{WT} PCI-13 cells EVs (left) and after transfection with miR-34a mimic and incubation with EVs from p53^{null} PCI-13 cells (right); data independently replicated in 12 wells. **b**, Quantification of neuritogenesis 72 h after neuron–EV co-culture. EV-free medium (grey bars) or non-specific antagomiR/miR-mimic transfection were used as controls ($n = 3$ biologically independent samples per condition). **c**, Quantification of neuritogenesis in TG neurons 72 h after neuron–EV co-culture using EVs from p53^{WT} OCSCC cells stably expressing a short hairpin targeting miR-34a (shmiR-34a) or non-specific control ($n = 4$ biologically independent samples per condition). **d**, Representative images showing β3-tubulin-positive neural fibres in orthotopic OCSCC xenografts stably expressing shmiR-34a ($n = 8$ mice) or non-specific control ($n = 7$ mice); data independently replicated in 15 mice. **e**, Quantification of neural density 3 weeks after orthotopic injection as in **d**. Bar graphs represent mean \pm s.e.m. Unpaired two-tailed t -test (**c**, **e**) or one-way ANOVA with Tukey multiple comparisons (**b**).

from p53^{null} OCSCC cells, significantly decreased the number of neurofilaments (Fig. 2a, b, Extended Data Fig. 3h). Co-culture of TG neurons with EVs derived from miR-34a-deficient p53^{WT} OCSCC cells¹³ (Extended Data Fig. 3i–k) significantly increased the number of neurofilaments in vitro compared with no EVs or EVs derived from the same cells without knockdown (Fig. 2c). Knockdown of miR-34a in p53^{WT} OCSCC cells also increased neuritogenesis in vivo, compared with the same cells without knockdown (Fig. 2d, e). Together, these data support the notion that EV shuttling of OCSCC-derived miR-34a to cancer-associated neurons negatively regulates, and confers resistance to, EV-derived axonogenic signals.

To dissect the relative contributions to axonogenesis of stimulatory and inhibitory signals in cancer-derived EVs, we used antagomiRs to selectively inhibit 13 miRNAs detected by our EV miRNA sequencing that have previously been shown to induce neural growth or cancerous neural invasion^{14–17}. Inhibition of miR-21, miR-197, or miR-324, but not of the other miRNAs, reduced neurite outgrowth in TG neurons co-cultured with EVs derived from miR-34a-deficient p53^{null} OCSCCs (Extended Data Fig. 4a). Transfection of TG neurons with miR-21, miR-197, or miR-324 alone only modestly increased neuritogenesis; however, co-transfection with both miR-21 and miR-324, with or without miR-197, increased neuritogenesis twofold (Extended Data Fig. 4b, c). Incubation of TG neurons with liposomes containing miR-21, miR-324, and scramble miRNA increased axonogenesis compared with liposomes containing miR-34a, miR-21 and miR-324 together (Extended Data Fig. 4d–f). These data are consistent with aberrant, but orchestrated, signalling by miRNAs in OCSCC-derived EVs that both negatively regulates and acts as a ligand that modulates the tumour neural microenvironment.

EVs drive sensory nerve reprogramming

To evaluate the effect of OCSCC-derived EVs on neuronal transcriptional programs, we performed RNA sequencing of human DRG neurons co-cultured with EVs from p53^{null} or p53^{WT} OCSCC cells¹⁸. Principal component analysis of RNA sequencing data showed that neurons cultured with EVs from p53^{WT} OCSCC cells were segregated from those cultured with EVs from p53^{null} OCSCC cells (Fig. 3a). Ingenuity pathway analysis revealed that the differentially expressed genes in the latter group of neurons were enriched for neuronal outgrowth and morphogenesis, synaptogenesis, differentiation and stemness, and synaptic transmission (Fig. 3b). To evaluate the effect of p53 on cancer-associated neuronal differentiation, we analysed nerve fibre densities in samples of tissue removed during glossectomy ($TP53^{+/+}$, $n = 12$; $TP53^{mut}$, $n = 12$) from patients with OCSCC treated at MD Anderson Cancer Center. Assessment of the sympathetic and parasympathetic branches of the autonomic nervous system showed that fibres positive for tyrosine hydroxylase (TH, adrenergic), but not those positive for vesicular acetylcholine transporter (parasympathetic), were significantly denser in $TP53^{mut}$ OCSCCs than in $TP53^{WT}$ tumours (Fig. 3c, d, Extended Data Fig. 5a–e). Similarly, in vivo, TH⁺ nerve fibres were significantly denser in tongue tumour specimens from $Trp53^{-/-}$ tumours excised from $Krt5^{Cre}Trp53^{flax/flax}$ mice than in those from $Trp53^{flax/flax}$ tumours (wild-type $Trp53$) excised from $Trp53^{flax/flax}$ control mice (Fig. 3e). Together these results suggest that the recruitment of proximal adrenergic neurons is related to signals that originate from p53-deficient tumours.

Next, we investigated whether shuttling of tumour-derived miRNAs in EVs regulates the differentiation of cancer-associated neurons. Neuritogenesis and noradrenaline release were increased in human DRG or mouse TG sensory neurons 72 h after incubation with EVs derived from p53^{null} OCSCC cells, but not from p53^{WT} OCSCC cells (Fig. 3f–h, Extended Data Fig. 5f–k). Next, we tested whether the adrenergic neurogenic response of cancer-associated neurons to epithelial loss of p53 could be rescued by EVs derived from p53^{WT} OCSCC cells in vivo. Daily intratumoural injections of EVs from p53^{WT} OCSCC cells markedly inhibited noradrenaline secretion and TH expression in orthotopic xenografts of p53^{null} OCSCC cells, compared with tumours injected with p53^{null} EVs and controls (Fig. 3i–k, Extended Data Fig. 5l). These results provide evidence that EVs derived from p53^{WT} OCSCC cells suppress neo-adrenergic cancer-associated neurogenesis.

miR-34a restricts cell fate and impedes somatic cell reprogramming, whereas miR-34a deficiency expands cell developmental potential^{19–22}. We tested the ability of miR-34a to suppress the tumour-associated phenotypic switch. Both TH expression and noradrenaline levels (Fig. 3l, m) were higher in TG neurons incubated with EVs purified from p53^{WT} OCSCC cells in which miR-34a was knocked down than in control TG neurons incubated with EVs from p53^{WT} OCSCC cells in which miR-34a was not knocked down. Furthermore, incubation of TG neurons with liposomes containing miR-21, miR-324, and scramble miRNA, but not with liposomes containing miR-21, miR-324, and miR-34a, resulted in robust noradrenaline synthesis (Extended Data Fig. 5m).

Next we investigated the transcriptomes of p53-deficient cancer-associated neurons. We identified 2,495 genes that were upregulated and 1,760 that were downregulated in neurons incubated with EVs from p53^{null} OCSCC cells, compared with neurons incubated with EVs from p53^{WT} OCSCC cells (Extended Data Fig. 5n; GEO accession number GSE140189). Upregulated genes were associated with neuronal survival, development, growth, and branching; downregulated genes were associated with neuronal function and synaptic transmission (Gene Ontology terms)²³. Neurons incubated with EVs from p53^{null} OCSCC cells had high expression of catecholamine biosynthesis-related genes (Extended Data Fig. 6a) and low expression of endogenous sensory neuron pain signalling genes including *Ntrk2*, *Tac1* and *Plcg1*; potassium channel family genes; and glutamate metabotropic receptor genes (*Prkc* and *Pka*). These analyses suggest that the TG sensory neurons

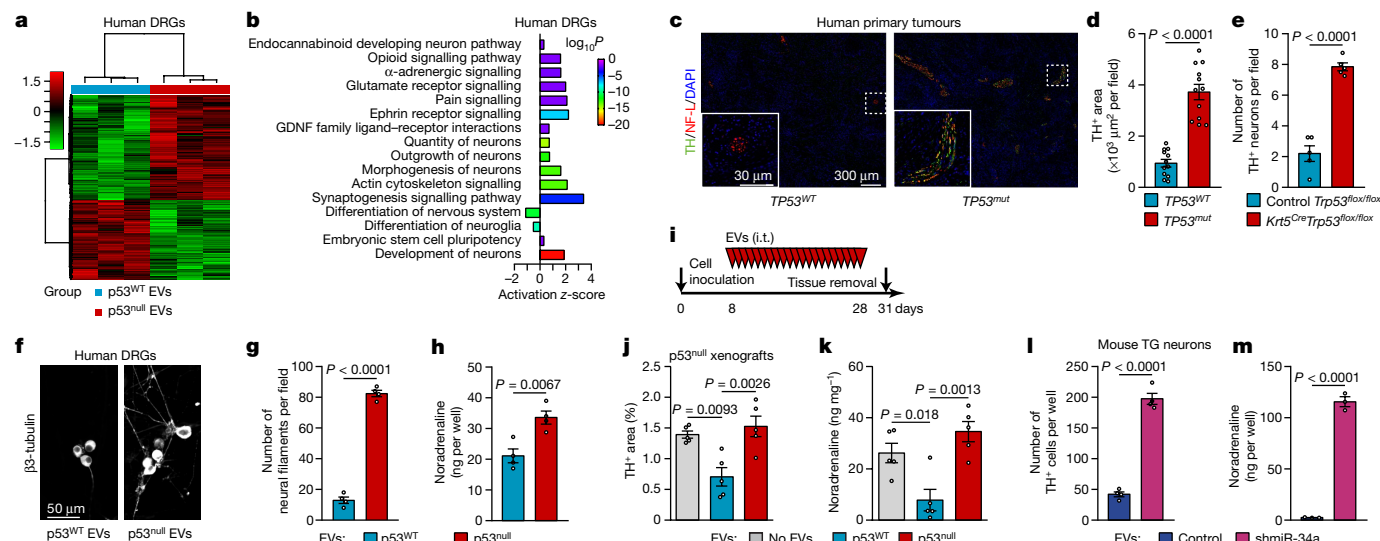


Fig. 3 | p53-deficient tumours are enriched with adrenergic nerve fibres.

a, b, Heat map (**a**) and enriched Gene Ontology terms (**b**) for differentially expressed genes in freshly collected human DRG neurons incubated with EVs from p53^{WT} or p53^{null} PCI-13 cells, plotted by fold enrichment with the associated log₁₀ P value (Fisher's exact algorithm for functional gene set enrichment); $n = 3$ biologically independent samples per condition. **c**, Representative images showing TH⁺ adrenergic neural fibres in TP53^{WT} or TP53^{null} human OSCC tissue; data independently replicated in 24 patient specimens. **d**, Quantification of TH⁺ areas as in **c** ($n = 12$ independent samples per group). **e**, Quantification of TH⁺ neural fibres in tumours from Krt5^{Cre}Trp53^{flax/flax} and control mice ($n = 5$). **f**, Representative images of neurites (β3-tubulin) in human DRG co-cultured with EVs from p53^{WT} or p53^{null}

PCI-13 cells; data independently replicated in 8 wells. **g, h**, Quantification of neuritogenesis (**g**) and noradrenaline levels (**h**) in human DRGs as in **f** ($n = 4$ biologically independent samples per condition). **i–k**, PCI-13-p53^{null} or thortopic tumours were injected daily with no EVs (vehicle) or with EVs from p53^{null} or p53^{WT} PCI-13 cells for 3 weeks ($n = 5$ mice per group, **i**), and TH⁺ neural areas (**j**) and noradrenaline levels (**k**) in the tongue were measured. **l, m**, TG neurons were co-cultured with EVs from p53^{WT} OSCC cells treated with lentiviral miR-34a or non-specific inhibitors. Quantification of TH⁺ TG neurons ($n = 4$ mice per condition, **l**), and noradrenaline levels ($n = 3$ biologically independent samples per condition, **m**). Bars represent mean ± s.e.m. Unpaired two-tailed *t*-test (**d, e, g, h, l, m**) or one-way ANOVA with Tukey multiple comparisons (**j, k**).

acquired transcriptional programs similar to those of sympathetic neurons.

To better understand the functional annotations of the involved miRNAs, we compared the transcriptional profiles of TG neurons transfected with miR-21, miR-34a, or miR-324 with that of TG neurons transfected with scramble miRNAs (Extended Data Fig. 6b). Neural identity determination domains (Gene Ontology terms) were significantly enriched in transcription factors crucial for neural cell fate and catecholaminergic differentiation, including *En1*, *Lrp6*, *Ryk*, *Shh*, *Fzd3*, *Erbp2*, and *Wnt5a* (induced by miR-21), and *Nrp2*, *Gdnf*, and *Sema3f* (induced by miR-324). Two determinants of sensory neuron differentiation, *Ntrk1* and *Isl1*, were downregulated by miR-21 and miR-324, respectively. Expression of genes involved in fundamental inhibitory neuron differentiation pathways, including *Eya1*, *Sox9*, *Homez*, *Cic*, *Kdm5a*, and *Eif4ebp3*, was dysregulated by miR-34a. Neural growth domains were enriched in genes associated with neurogenesis, axonogenesis, and neuron projection development pathways. The axon guidance genes, most of which were deregulated by miR-21, included members of the neural cell adhesion family (*Cdh2*, *Nrcam*, *Lamc1*, and *Nfasc*), semaphorins (*Plxna1*, *Plxna2*, *Sema3b*, *Sema3d*, and *Sema3e*), ephrins (*Efnb2*, *Epha3*, *Ephb4*, and *Epha4*), Rho GTPases (*Srgap2*), and *Napa*, *Slit3*, and *Slitrk6*. The top pathways annotated in the neuron function domain were mostly related to synaptic functions, ion channels, and neurotransmission. Glutamatergic signalling genes (for example, *Cnih2*, *Ntrk2*, *Oprm1*, and *Syt1*) and nociception genes (*Ptgs1*, *Tac1*, and *Npy1r*) were also downregulated by miR-324, whereas key components of catecholaminergic neuron maintenance, including *Nr4a2*, *En1*, *Lrp6*, and *Ryk*, were upregulated by miR-21.

To investigate the origin of the TH⁺ neo-nerves that infiltrated OSCC xenografts, we ablated adrenergic nerves by injecting mice with 6-hydroxydopamine (6-OHDA) before inoculating them with tumour cells (Fig. 4a). Quantification of TH⁺ fibres showed that there

were similar increases in the numbers of cancer-associated TH⁺ fibres in both 6-OHDA-treated mice and controls (Fig. 4b, c). In 6-OHDA-treated mice, the levels of noradrenaline were significantly higher in p53^{null} tumours than in p53^{WT} tumours ($P < 0.001$ for each, pairwise comparisons, Fig. 4d). Flow cytometry analysis of TG neurons in 6-OHDA-treated mice showed a substantial increase in TH⁺ neurons after implantation of p53^{null} tumour cells (Extended Data Fig. 7a). To investigate whether differentiation signals modify the neural signalling components in the tumour microenvironment, we surgically cut the lingual nerve, a branch of the trigeminal nerve, thereby ablating sensory innervation to the tongue, before orthotopic injection of tumour cells into BALB/c (*nu/nu*) mice. Surgical lingual denervation markedly decreased the cancer-associated TH⁺ area ($P < 0.0001$, Fig. 4e, f), the number of TH⁺ ipsilateral TG neurons (Extended Data Fig. 7b, c), and tumour development (Fig. 4g) relative to those in sham-operated mice. However, chemical sympathectomy before orthotopic xenograft implantation of OSCC cells did not affect tumour growth (Fig. 4h). These results indicate that tumour-derived signals regulate the adrenergic differentiation of cancer-associated nerves and that these neo-adrenergic nerves, rather than infiltration of pre-existing adrenergic nerves, promote tumour growth.

In nude mice orthotopically xenografted with p53^{null} OSCC cells, the number of TH⁺ cell bodies was markedly increased compared with tumour-free mice or mice orthotopically xenografted with p53^{WT} OSCC cells (Extended Data Fig. 7d, e). Furthermore, inoculation of mice with p53^{WT} OSCC cells in which miR-34a was knocked down increased TH expression and noradrenaline secretion; these increases, along with tumour volume, were abrogated by lingual denervation (Extended Data Fig. 7f–i). We further explored the cancer-induced retrograde reprogramming of TG neurons in vivo. We identified 685 genes that were upregulated and 15 that were downregulated in TG neurons from mice orthotopically injected with p53^{null} OSCC cells,

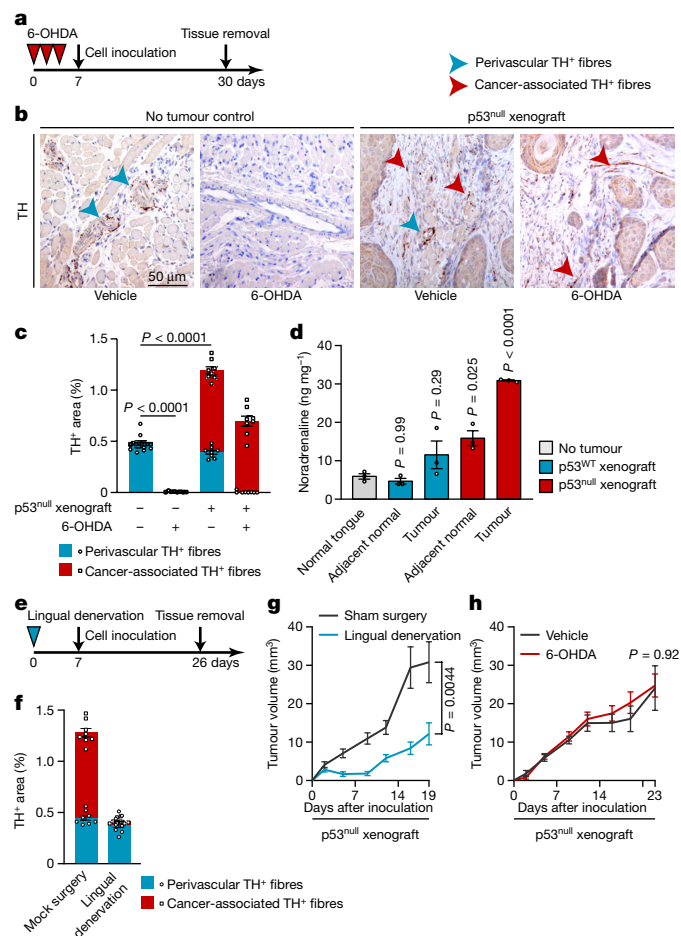


Fig. 4 | De novo transdifferentiated cancer-associated adrenergic nerves support tumour growth. **a**, BALB/c (*nu/nu*) mice were chemically sympathectomized by intraperitoneal injection of 6-OHDA and then injected orthotopically with p53^{null} PCI-13 cells. Tumour volume was monitored for 3 weeks. **b**, Illustrative immunohistochemical analysis for TH⁺ neural fibres in tongues from mice with or without tumours and with or without sympathectomy. Blue arrowheads, pre-existing TH⁺ perivascular neural fibres; red arrowheads, non-perivascular TH⁺ fibres that emerged after tumour formation (cancer-associated nerves); data independently replicated in 32 mice. **c**, Perivascular and cancer-associated TH⁺ area ($n = 8$ mice per condition). **d**, Noradrenaline levels in tumour-bearing tongue (ipsilateral to tumour injection site), adjacent normal tongue (contralateral to tumour injection site), and normal tongue controls ($n = 3$). **e**, BALB/c (*nu/nu*) mice were orthotopically xenografted with p53^{null} PCI-13 cells to the denervated or sham-operated tongue. **f**, Quantification of TH⁺ nerve fibres in mice from **e** ($n = 8$ mice per condition). **g**, **h**, In vivo p53^{null} PCI-13 tumour growth after lingual denervation (blue, $n = 12$) compared with sham surgery controls ($n = 10$, **g**), or chemical sympathectomy (red, $n = 12$) compared with vehicle controls ($n = 11$, **h**). Bar graphs and tumour growth curves represent mean \pm s.e.m. Unpaired two-tailed *t*-test (**g**, **h**) or one-way ANOVA with Tukey multiple comparisons (**c**, **d**, **f**).

compared with mice injected with p53^{WT} OCSCC cells (Extended Data Fig. 8a). Ingenuity pathway analysis indicated significant activation of embryonic stem cell pluripotency canonical pathways (Extended Data Fig. 8a). We also found statistically significant activation of adrenergic signalling pathways, axonogenesis, neurite branching, and ephrin axonal guidance signalling.

To elucidate the reprogramming of neuronal identity^{24,25}, we profiled the expression of known neuronal lineage differentiation transcription factors (Fig. 5a, Extended Data Fig. 8b–e) in mouse TG after orthotopic inoculation of mice with p53-deficient or p53-sufficient OCSCC cell lines. POU5F1 and KLF4 not only have essential roles in differentiation

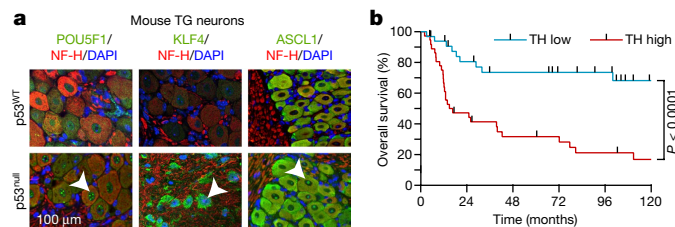


Fig. 5 | Retrograde signalling by p53-deficient but not p53-sufficient cancer cells activates neural reprogramming. **a**, Representative images demonstrating expression of POU5F1, KLF4, and ASCL1 in ipsilateral TGs after orthotopic injection of p53^{WT} (upper) or p53^{null} (lower) PCI-13 cells; data independently replicated in six mice. **b**, Kaplan–Meier curves showing the overall survival of patients with high and low TH⁺ adrenergic nerve densities. Two-sided log-rank test.

but also are validated direct targets of miR-34a²⁶, alongside strong candidate targets such as NEUROG2 and ASCL1, which are crucial for the determination of neuronal identity^{27,28}. The transcriptional factors POU5F1 and KLF4, which are sufficient to reprogram mouse adult neural stem cells^{24,29,30}, showed increased expression in mouse TG neurons after inoculation with p53^{null} OCSCC cells (Fig. 5a). In line with the neonuron adrenergic phenotype, the number of neurons expressing ASCL1, which is essential for proper development of the sympathetic nervous system^{24,31–33}, was elevated in mouse TG after inoculation with p53^{null} but not p53^{WT} OCSCC cells (Fig. 5a).

Adrenergic innervation promotes OCSCC growth

To evaluate the effect of selective adrenergic receptor blockade on OCSCC progression, we orthotopically injected p53-deficient OCSCC cells into BALB/c (*nu/nu*) mice treated with carvedilol, a non-selective blocker of β_1 , β_2 , and α_1 adrenergic receptors. Tumours from carvedilol-treated mice exhibited lower growth rates and proliferation (Ki-67⁺) indices than did tumours from vehicle-treated mice, with similar cardiovascular haemodynamics (Extended Data Fig. 8f–h, Supplementary Table 1). These data support the notion that adrenergic neuron-derived signals have an important role in OCSCC progression. In a validation set of patients treated at MD Anderson Cancer Center ($n = 70$, Supplementary Table 2), Kaplan–Meier analysis revealed that increased TH⁺ nerve density (Fig. 5b, Extended Data Fig. 8i) in OCSCCs was associated with lower recurrence-free survival rates ($P = 0.00103$, log-rank test) and lower overall survival rates ($P < 0.0001$, log-rank test). The statistical significance of the association with adrenergic nerve fibre densities was sustained in multivariable analysis after adjustment for clinical variables, including age, sex, pathologic stage, surgical margin status, perineural invasion presence, and treatment modalities (Supplementary Table 3). These results suggest that nerve density assessment merits exploration as an independent predictive marker of oral cancer aggressiveness.

Discussion

Neural regulation represents an emerging targetable pathway for the treatment of cancer^{1,34–36}. The peripheral adrenergic nervous system has previously been shown to regulate prostate cancer tumorigenesis^{1,3,34,35}. Sympathetic nerves form a dominant part of the normal prostate microenvironment, while in the oral cavity, their presence is modest and limited to the perivascular space. Our present study reveals that the emergence of adrenergic neonerves in the tumour microenvironment accompanies the initial phase of OCSCC development. In contrast to previous findings using a prostate cancer mouse model, in our OCSCC mouse model, ablation of the sympathetic nervous system

before tumour inoculation neither abrogated the development of adrenergic neo-nerves nor inhibited tumour growth¹.

Neither the origin of adrenergic neo-nerves nor the cellular and molecular events that control their development throughout tumorigenesis have previously been characterized. A recent study did show that doublecortin-expressing neural progenitors from the central nervous system infiltrate prostate tumours and metastases³³. Here, we have identified crosstalk between the peripheral nervous system and head and neck tumours and described a phenotypic switch, induced by cancer cells, in which sensory nerves differentiate into adrenergic neo-neurons. Our findings show that in p53-deficient tumours, an miRNA-based mechanism mediates neuronal responses to environmental cues and determines the fate of cancer-associated neurons. We have shown that axonal sprouting and autonomic reprogramming of existing nerves occur as a result of miRNA shuttling from cancer cells to neurons. These miRNAs orchestrate gene expression via combined dominantly negative (for example, miR-34a) and positive (for example, miR-21 and miR-324) effects, activating transcriptional programs that establish neuronal identity. In our mouse model of OCS, surgical ablation of sensory nerves prevented the development of these adrenergic neo-nerves. Our results thus show that the peripheral sensory nerves may be reprogrammed during the development of cancer in a manner similar to that of neural progenitors that initiate adrenergic neurogenesis during tumour formation.

As tumours evolve, neo-neural networks develop in and around the tumour stroma, providing signals that coordinate cancer progression^{1,2}. These results are consistent with recent preclinical data suggesting that sympathetic fibres accumulate in the normal vicinity of solid tumour tissues and infiltrate into the stroma. Furthermore, clinical data show that cancer patients treated with β -blockade have improved survival, supporting the role of adrenergic nerve activity in cancer progression^{35,37}. Although further studies will be required to dissect the molecular events that link tumour-associated neuritogenesis to cancer progression, our data raise the tantalizing possibility that drugs that target both axonal growth and the adrenergic nervous system could be useful for the treatment of head and neck cancer.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-1996-3>.

- Magnon, C. et al. Autonomic nerve development contributes to prostate cancer progression. *Science* **341**, 1236361 (2013).
- Ayala, G. E. et al. Cancer-related axonogenesis and neurogenesis in prostate cancer. *Clin. Cancer Res.* **14**, 7593–7603 (2008).
- Zahalka, A. H. et al. Adrenergic nerves activate an angio-metabolic switch in prostate cancer. *Science* **358**, 321–326 (2017).
- Ravi, R. et al. Regulation of tumor angiogenesis by p53-induced degradation of hypoxia-inducible factor 1 α . *Genes Dev.* **14**, 34–44 (2000).
- Schmid, J. O. et al. Cancer cells cue the p53 response of cancer-associated fibroblasts to cisplatin. *Cancer Res.* **72**, 5824–5832 (2012).
- Antoniades, H. N., Galanopoulos, T., Neville-Golden, J., Kiritsy, C. P. & Lynch, S. E. p53 expression during normal tissue regeneration in response to acute cutaneous injury in swine. *J. Clin. Invest.* **93**, 2206–2214 (1994).

- Yun, M. H., Gates, P. B. & Brockes, J. P. Regulation of p53 is critical for vertebrate limb regeneration. *Proc. Natl Acad. Sci. USA* **110**, 17392–17397 (2013).
- Li, Z. et al. Cdkn2a suppresses metastasis in squamous cell carcinomas induced by the gain-of-function mutant p53(R172H). *J. Pathol.* **240**, 224–234 (2016).
- Joerger, A. C. & Fersht, A. R. Structural biology of the tumor suppressor p53. *Annu. Rev. Biochem.* **77**, 557–582 (2008).
- Ostrowski, M. et al. Rab27a and Rab27b control different steps of the exosome secretion pathway. *Nat. Cell Biol.* **12**, 19–30 (2010).
- Posner, R. et al. Neuronal small RNAs control behavior transgenerationally. *Cell* **177**, 1814–1826.e1815 (2019).
- Chen, W. & Qin, C. General hallmarks of microRNAs in brain evolution and development. *RNA Biol.* **12**, 701–708 (2015).
- Pang, R. T. et al. MicroRNA-34a suppresses invasion through downregulation of Notch1 and Jagged1 in cervical carcinoma and choriocarcinoma cells. *Carcinogenesis* **31**, 1037–1044 (2010).
- Prueitt, R. L. et al. Expression of microRNAs and protein-coding genes associated with perineural invasion in prostate cancer. *Prostate* **68**, 1152–1164 (2008).
- Lubov, J. et al. Meta-analysis of microRNAs expression in head and neck cancer: uncovering association with outcome and mechanisms. *Oncotarget* **8**, 55511–55524 (2017).
- Ramdas, L. et al. miRNA expression profiles in head and neck squamous cell carcinoma and adjacent normal tissue. *Head Neck* **31**, 642–654 (2009).
- Sousa, L. O. et al. Lymph node or perineural invasion is associated with low miR-15a, miR-34c and miR-199b levels in head and neck squamous cell carcinoma. *BBA Clin.* **6**, 159–164 (2016).
- North, R. Y. et al. Electrophysiological and transcriptomic correlates of neuropathic pain in human dorsal root ganglion neurons. *Brain* **142**, 1215–1226 (2019).
- Choi, Y. J. et al. miR-34 miRNAs provide a barrier for somatic cell reprogramming. *Nat. Cell Biol.* **13**, 1353–1360 (2011).
- Choi, Y. J. et al. Deficiency of microRNA miR-34a expands cell fate potential in pluripotent stem cells. *Science* **355**, eaag1927 (2017).
- Agostini, M. et al. Neuronal differentiation by TAp73 is mediated by microRNA-34a regulation of synaptic protein targets. *Proc. Natl Acad. Sci. USA* **108**, 21093–21098 (2011).
- Jauhari, A., Singh, T., Singh, P., Parmar, D. & Yadav, S. Regulation of miR-34 family in neuronal development. *Mol. Neurobiol.* **55**, 936–945 (2018).
- Dennis, G. Jr et al. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol.* **4**, 3 (2003).
- Tsunemoto, R. et al. Diverse reprogramming codes for neuronal identity. *Nature* **557**, 375–380 (2018).
- Vierbuchen, T. et al. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**, 1035–1041 (2010).
- Hsu, S. D. et al. miRTarBase update 2014: an information resource for experimentally validated miRNA-target interactions. *Nucleic Acids Res.* **42**, D78–D85 (2014).
- Miranda, K. C. et al. A pattern-based method for the identification of microRNA binding sites and their corresponding heteroduplexes. *Cell* **126**, 1203–1217 (2006).
- Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P. & Burge, C. B. Prediction of mammalian microRNA targets. *Cell* **115**, 787–798 (2003).
- Kim, J. et al. Direct reprogramming of mouse fibroblasts to neural progenitors. *Proc. Natl Acad. Sci. USA* **108**, 7838–7843 (2011).
- Yao, H. et al. Transdifferentiation-induced neural stem cells promote recovery of middle cerebral artery stroke rats. *PLoS One* **10**, e0137211 (2015).
- Wapinski, O. L. et al. Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. *Cell* **155**, 621–635 (2013).
- Axelsson, H. The Notch signaling cascade in neuroblastoma: role of the basic helix-loop-helix proteins HASH-1 and HES-1. *Cancer Lett.* **204**, 171–178 (2004).
- Mauffrey, P. et al. Progenitors from the central nervous system drive neurogenesis in cancer. *Nature* **569**, 672–678 (2019).
- Zhao, C. M. et al. Denervation suppresses gastric tumorigenesis. *Sci. Transl. Med.* **6**, 250ra115 (2014).
- Melhem-Bertrandt, A. et al. Beta-blocker use is associated with improved relapse-free survival in patients with triple-negative breast cancer. *J. Clin. Oncol.* **29**, 2645–2652 (2011).
- Faulkner, S., Jobling, P., March, B., Jiang, C. C. & Hondermarck, H. Tumor neurobiology and the war of nerves in cancer. *Cancer Discov.* **9**, 702–710 (2019).
- Lin, C. S., Lin, W. S., Lin, C. L. & Kao, C. H. Carvedilol use is associated with reduced cancer risk: a nationwide population-based cohort study. *Int. J. Cardiol.* **184**, 9–13 (2015).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Methods

Animals and in vivo procedures

B6.129P2-Trp53^{tm1BrnJ} (*Trp53*^{flax/flax}) and BALB/c nu/nu (B6.Cg-*Foxn1*^{nu+/-}) mice were obtained from The Jackson Laboratory. *Krt5*^{cre} mice were obtained from Dr. Carlos Caulin³⁸. All animal studies were carried out according to protocols approved by The University of Texas MD Anderson Cancer Center Institutional Animal Care and Use Committee. Mouse housing, husbandry, and care practices met or exceeded the minimum requirements set forth in the Animal Welfare Act and the Guide for the Care and Use of Laboratory Animals (8th Edition). Disease development and progression were closely monitored, and mice with metastases were euthanized as soon as we noticed signs of discomfort in the mice or when the largest dimension of a tumour reached 5 mm, according to our approved protocol. In none of the experiments were these limits exceeded.

Orthotopic human tumours were implanted by injection of 5×10^4 PCI-13 cells suspended in 30 μ l serum-free medium (see Supplementary Methods) into the lateral tongue of 6- to 8-week-old BALB/c nu/nu mice and were monitored three times per week as previously described³⁹. For the transgenic models, we crossbred *Trp53*^{flax/flax} and *Krt5*^{cre} mice to generate male and female *Krt5*^{cre}*Trp53*^{flax/flax} mice; the recombined epithelial tissue of these mice lacks p53. We added 4NQO (100 μ g/ml) to the drinking water (1% sucrose) of *Krt5*^{cre}*Trp53*^{flax/flax} mice and *Trp53*^{flax/flax} controls^{38,40,41}. After 8 weeks of 4NQO treatment, mice were killed at 0, 20, and 30 weeks to study normal, precursor, and malignant lesion innervation, respectively. All animals underwent a full oral cavity examination three times per week and were killed for tissue retrieval 4 weeks after tumour cell injection.

For the assessment of in vivo neural recruitment, mice were anaesthetized and prepared: the right chorda-lingual nerve was exposed in the neck and transected between the anterior belly of the digastric and masseter muscles⁴². Although the proximal and distal stumps of the transected chorda-lingual nerve were separated, we resected a 5-mm section of each stump to minimize regeneration.

CRISPR-Cas9 knockout

For generation of RAB27A and RAB27B knockout cells, two synthetic single-guide RNAs (sgRNAs) in complex with Cas9, targeting the protein coding sequence of either *RAB27A* (guide sequence: GTC GTT AAG CTA CGA AAC CT, exon 5) or *RAB27B* (guide sequence: TGA ACC GCA AGC TCG GGA AC, exon 5), were transfected by electroporation. sgRNAs and Cas9 2NLS Nuclease were purchased from Synthego, and electroporation was performed using the Cell Line Nucleofector Kit V (Lonza). A total of 100,000 cells were diluted in 50 μ l electroporation buffer containing 3.6 μ M of each sgRNA and 0.8 μ M Cas9 enzyme, and electroporated using the program P-020 (Amaxa Nucleofector II). Cells were transferred to full medium immediately after electroporation and left to recover for 1 week, and then single-cell colonies were generated by seeding of a single cell using flow cytometry (sorting for live cells by propidium iodide staining) at the South Campus Flow Cytometry and Cellular Imaging Core Facility at MD Anderson Cancer Center. Screening for knockout clones was done by Sanger sequencing of target regions and western blotting.

Isolation, quantification, and characterization of EVs

Conditioned media were collected, and EVs were isolated by differential centrifugation and analysed using NanoSight, as previously described⁴³. In brief, conditioned media were centrifuged at 300g for 10 min to eliminate cells and at 2,000g and 10,000g to eliminate dead cells and cell debris, respectively. Then, EVs were pelleted by ultracentrifugation at 120,000g for 70 min and subsequently washed with PBS at a similar speed. The number and size of EVs were determined as previously described using NanoSight analysis⁴³. In brief, EVs were analysed using a NanoSight LM10 Nanoparticle Characterization

system. All nanoparticle tracking analyses were carried out with identical experiment settings. Particles were measured for 60 s, and for optimal results, microvesicle concentrations were adjusted to obtain about 50 microvesicles per field of view. EV morphology was assessed by transmission electron microscopy as previously described⁴⁴.

For transmission electron microscopy, samples were placed on 100 mesh carbon-coated, formvar-coated copper grids treated with poly-L-lysine for approximately 1 h. Samples were then negatively stained with Millipore-filtered aqueous 1% uranyl acetate for 1 min. Stains were blotted dry from the grids with filter paper, and samples were allowed to dry. Samples were then examined in a JEM1010 transmission electron microscope (JEOL, USA, Inc.) at an accelerating voltage of 80 Kv. Digital images were obtained using the AMT Imaging System (Advanced Microscopy Techniques Corp.).

EVs were lysed in a 2% sodium dodecyl sulphate buffer, and equal amounts of protein were loaded onto a sodium dodecyl sulphate-polyacrylamide gel and transferred onto polyvinylidene difluoride membranes (Bio-Rad Laboratories). Antibodies against CD63 were used as primary antibodies. As secondary antibodies, horseradish peroxidase-linked antibodies against rabbit immunoglobulin G (GE Healthcare) were used at a dilution of 1:5,000. Bound antibodies were visualized by chemiluminescence.

Neuron mRNA sequencing and data analysis

Low-input RNA libraries compatible with Illumina were prepared using the Smart-Seq V4 Ultra Low Input RNA (Clontech) and KAPA HyperPlus Library Preparation kits. In brief, full-length, double-stranded cDNA was generated from 10 ng total RNA using Clontech's SMART (Switching Mechanism at 5' End of RNA Template) technology. The full-length double-stranded cDNA was amplified by eight cycles of long-distance PCR, then purified using AMPure Beads (Agencourt). Following bead elution, the cDNA was evaluated for size distribution and quantity using the 4200 TapeStation High Sensitivity DNA Kit (Agilent Technologies) and the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific), respectively. The cDNA was enzymatically fragmented, and 5 ng of the fragmented cDNA was used to generate Illumina-compatible libraries using the KAPA HyperPlus Library Preparation kit. The KAPA libraries were purified and enriched with eight cycles of PCR to create the final cDNA library. The libraries were quantified using the Qubit dsDNA HS Assay (Thermo Fisher Scientific), then multiplexed with 12–16 libraries per pool. The pooled libraries were quantified by quantitative PCR using the KAPA Library Quantification Kit (KAPA Biosystems) and assessed for size distribution using the 4200 TapeStation (Agilent Technologies). The libraries were then sequenced, one pool per lane, on the Illumina HiSeq4000 sequencer using the 76-bp paired-end format.

Paired-end reads in FASTQ format were initially checked for read quality using FastQC⁴⁵ and then aligned to the reference genome, UCSC mouse mm10 or GENCODE human GRCh38, using TopHat2⁴⁶. Alignment quality was evaluated from the output of TopHat2. The BAM file with mapped reads for each sample was sorted using SAMtools⁴⁷, which serves as an input for HTSeq⁴⁸ to estimate the number of reads that are mapped to each gene. The read counts for all samples were then normalized using the trimmed mean of M method implemented in the R Bioconductor package edgeR⁴⁹. Weakly expressed genes were excluded if they did not have more than one read per million in at least two samples (adjusting for library size for each sample). Principal component analysis and hierarchical cluster analysis using Pearson distance and Ward's minimum variance method were used to evaluate sample quality and similarity. The generalized linear model likelihood ratio test implemented in the edgeR package⁵⁰ was applied to determine significant differentially expressed genes between groups. Benjamini-Hochberg correction was applied to the *P* values for multiple testing adjustment. Significant differentially expressed genes were selected using a false discovery rate cutoff of 5% with or without an absolute fold-change of ≥ 2 for heatmap generation. The Pearson distance and

Ward's minimum variance method were used to generate the cluster dendrograms on the heatmaps.

EV miRNA sequencing and data analysis

miRNA libraries compatible with Illumina were prepared using the QIAseq miRNA Library Kit (QIAGEN), per the manufacturer's protocol. In brief, 80 ng of total RNA was sequentially ligated to adapters. In the first ligation step, an adenylated 3' DNA adaptor was ligated to the 3' end of the miRNA. In the second step of the ligation, an RNA adaptor was ligated to the 5' end of the mature miRNAs in the sample. A cDNA library was then synthesized from the mature miRNAs using a reverse transcription primer containing an integrated unique molecular barcode. During reverse transcription, the reverse transcription primer hybridized to the 3' adaptor and converted the dual 3'/5' ligated miRNAs to cDNAs, while adding a unique molecular barcode and a universal sample index to every miRNA molecule. Following bead cleanup, the libraries were enriched and unique sample indexes were added using 16 cycles of PCR. A post-PCR bead cleanup was performed, and then the libraries were assessed for size distribution using the TapeStation (Agilent Technologies), quantified using the Qubit assay (Thermo Fisher Scientific), and then multiplexed, 33 samples per pool. The pooled library was quantified by quantitative PCR using the KAPA Library Quantification Kit (KAPA Biosystems), then sequenced, one pool per run, on the Illumina NextSeq 500 Sequencer using the 75-nt SR High-Output flow cell.

Single-end reads in FASTQ format were initially checked for read quality using FastQC⁴⁵. The QIAseq miRNA Library Kit specific 3' adaptor sequence was trimmed from the 3' ends of sequencing reads using cutadapt⁵¹. The trimmed reads were then aligned to the reference genome, human GRCh38, using BWA⁵². SAMtools flagstat was used to check the mapping quality⁴⁷. The SAM files with mapped reads for each sample were sorted by coordinate and outputted in BAM format using the Picard tool of SortSam (<http://broadinstitute.github.io/picard/>). The miRNA GFF annotation file was downloaded from miRBase, which serves as an input for featureCounts⁵³ to estimate the number of reads that are mapped to each mature miRNA. Weakly expressed miRNAs were filtered out if they did not have more than one read in at least two samples. The read counts for all samples were then normalized using the trimmed mean of M method implemented in the R Bioconductor package edgeR⁴⁹. Principal component analysis and hierarchical cluster analysis using Pearson distance and Ward's minimum variance method were used to evaluate sample quality and similarity. The generalized linear model likelihood ratio test implemented in the edgeR package⁵⁰ was applied to determine significant differentially expressed miRNAs between groups. The Benjamini–Hochberg correction was applied to the *P* values for multiple testing adjustment. Significant differentially expressed miRNAs were selected using a false discovery rate cutoff of 5% for each comparison. The Pearson distance and Ward's minimum variance method were used to generate the cluster dendrograms on the heatmaps.

Affymetrix gene expression microarray and analysis

Total RNA was collected to assess global gene changes after TG neurons were transfected with miR-21, miR-34a, miR-324, or scramble miRNA. The array was analysed using the Affymetrix Clariom S mouse assay. The CEL files generated were processed through Transcriptome Analysis Console version 4.0 (Thermo Fisher Scientific), which normalized (and applied the log₂ function to) array signals using a robust multiarray averaging algorithm. Differential expression between neurons transfected with different miRNAs was defined as a fold-change in the absolute value that was equal to or greater to 1.1 and a *P* value obtained from the moderated *t*-statistic from the limma package that was less than 0.05. The gene-level differential expression analysis was performed using Transcriptome Analysis Console version 4.0. Pathway analysis and functional annotation for

upregulated and downregulated genes in the three comparisons were performed using enrichR (<https://amp.pharm.mssm.edu/Enrichr/>). Dysregulated genes from the pathways of interest were displayed using R software (version 3.5.1).

Canonical pathway integrative analysis

Canonical pathway activation and downregulation were predicted in QIAGEN Ingenuity pathway analysis. Ingenuity pathway analysis was used to identify the cascade of upstream and downstream regulators of the core gene set. Ingenuity pathway analysis uses a priori knowledge of expected interactions between transcriptional regulators and their target genes stored in Ingenuity Knowledge Base, a scientific literature-based database <https://www.g6g-software-directory.com/bio/cross-omics/dbs-kbs/20018U-Ingenuity-Knowledge-Base.php>.

miRNA target analysis

Predicted targets were retrieved from TargetMiner⁵⁴, TargetScan⁵⁵, and miRDB^{56,57} databases through the searchable database miR-Base⁵⁸, while experimentally validated targets were retrieved from miRTarBase⁵⁹.

Preparation of miRNA encapsulated liposomes

A lipid mixture of 1,2-dipalmitoyl-3-dimethylammonium-propane (Avanti Polar Lipids), cholesterol (Sigma-Aldrich), and DSPE-PEG 2000 (Avanti Polar Lipids) in a molar ratio of 42:48:10 was dissolved in 100% ethanol. The mixed lipids were added to 125 mmol/l sodium acetate buffer (pH 5.2) to yield a solution containing 35% ethanol. The resultant nanoparticles were extruded through a 0.08-μm membrane (Whatman) using a LIPEX Extruder (Northern Lipids) to form 120- to 140-nm nanoparticles. miRNA in 50 mmol/l sodium acetate (pH 5.2) and 35% ethanol was added to the nanoparticles at 538:1 lipid:miRNA ratios and incubated at 37 °C for 30 min. Ethanol removal and buffer exchange of miRNA-containing nanoparticles were achieved by dialysis against PBS using a 13,000-kDa MWCO dialysis bag (Spectra/Por Dialysis Membrane) for 24 h at 4 °C, with the external medium exchanged after 1, 3, and 24 h. Finally, the formulation was filtered through a 0.2-μm sterile filter. Particle size and zeta potential were determined using a Zetasizer Nano ZS (Malvern). miRNA entrapment efficiency was determined by the Quant-iT RiboGreen RNA assay (Invitrogen). For labelling, 50 μl of 1 mg/ml 18:1 Liss Rhod PE 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine-N-(lissamine rhodamine B sulfonyl) (ammonium salt) was added to the lipid mixture.

Statistical analysis

The unpaired, two-tailed *t*-test and one-way ANOVA with Tukey multiple comparisons were carried out to analyse in vitro data. For mouse studies, a two-way ANOVA was used to compare tumour volumes between control and treatment groups. For immunohistochemical analyses, a one-way ANOVA was used to compare control and treatment groups. Survival was analysed by the Kaplan–Meier method and compared using the log-rank test. All bar graphs were expressed as mean ± s.e.m. with individual values shown if *n* ≤ 12. *P* values of less than 0.05 were considered to indicate nominal statistical significance. On the basis of the variance of xenograft growth in control mice, we used at least three mice per genotype to give 80% power to detect an effect size of 20% with a significance level of 0.05. For all mouse experiments, the number of independent mice used is listed in the figure legend. No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Neuron RNA sequencing data from in vivo and in vitro experiments are available from the Gene Expression Omnibus (GEO) under accession number GSE134220. mRNA array data are available on GEO under accession number GSE140189, and miRNA array data are available on GEO under accession number GSE140324. All other data are available in the article and source data, or from the corresponding author upon reasonable request.

38. Caulin, C. et al. An inducible mouse model for skin cancer reveals distinct roles for gain- and loss-of-function p53 mutations. *J. Clin. Invest.* **117**, 1893–1901 (2007).
39. Myers, J. N., Holsinger, F. C., Jasser, S. A., Bekele, B. N. & Fiddler, I. J. An orthotopic nude mouse model of oral tongue squamous cell carcinoma. *Clin. Cancer Res.* **8**, 293–298 (2002).
40. Foy, J. P. et al. The dynamics of gene expression changes in a mouse model of oral tumorigenesis may help refine prevention and treatment strategies in patients with oral cancer. *Oncotarget* **7**, 35932–35945 (2016).
41. Tang, X. H., Knudsen, B., Bemis, D., Tickoo, S. & Gudas, L. J. Oral cavity and esophageal carcinogenesis modeled in carcinogen-treated mice. *Clin. Cancer Res.* **10**, 301–313 (2004).
42. Guagliardo, N. A. & Hill, D. L. Fungiform taste bud degeneration in C57BL/6J mice following chorda-lingual nerve transection. *J. Comp. Neurol.* **504**, 206–216 (2007).
43. Ricklefs, F. et al. Extracellular vesicles from high-grade glioma exchange diverse pro-oncogenic signals that maintain intratumoral heterogeneity. *Cancer Res.* **76**, 2876–2881 (2016).
44. Haderk, F. et al. Extracellular vesicles in chronic lymphocytic leukemia. *Leuk. Lymphoma* **54**, 1826–1830 (2013).
45. FASTQC: a quality control tool for high throughput sequence data (Babraham Bioinformatics, 2011).
46. Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
47. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
48. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
49. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
50. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
51. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
52. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
53. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
54. Bandyopadhyay, S. & Mitra, R. TargetMiner: microRNA target prediction with systematic identification of tissue-specific negative examples. *Bioinformatics* **25**, 2625–2631 (2009).
55. Agarwal, V., Bell, G. W., Nam, J. W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *eLife* **4**, (2015).
56. Wong, N. & Wang, X. miRDB: an online resource for microRNA target prediction and functional annotations. *Nucleic Acids Res.* **43**, D146–D152 (2015).
57. Wang, X. Improving microRNA target prediction by modeling with unambiguously identified microRNA-target pairs from CLIP-ligation studies. *Bioinformatics* **32**, 1316–1322 (2016).
58. Kozomara, A. & Griffiths-Jones, S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* **42**, D68–D73 (2014).
59. Chou, C. H. et al. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res.* **46**, D296–D302 (2018).

Acknowledgements G.A.C. is the Felix L. Haas Endowed Professor in Basic Science. Work in the Calin laboratory is supported by National Institutes of Health (NIH/NCATS) grant UH3TR00943-01 through the NIH Common Fund, Office of Strategic Coordination (OSC), NCI grants 1R01CA182905-01 and 1R01CA222007-01A1, National Institute of General Medical Sciences (NIGMS) grant 1R01GM122775-01, U54 grant CA096297/CA096300 – UPR/MDACC Partnership for Excellence in Cancer Research 2016 Pilot Project, US Department of Defense grant CA160445P1, a Chronic Lymphocytic Leukemia Moonshot Flagship Project, a Sister Institution Network Fund (SINF) 2017 grant, and the Estate of C. G. Johnson Jr. Work in the Dougherty laboratory is supported by NIH grant CA200263, Thompson Family Foundation Initiative; P.M.D. is the H.E.B. Endowed Professor in Basic Science. The NIH Cancer Center Support Grant P30CA016672 supports the High Resolution Electron Microscopy Facility (K. Dunner Jr) and the Advanced Technology Genomics Core (core grant CA016672) at The University of Texas MD Anderson Cancer Center. We thank M. Sushnitha for assistance with miRNA encapsulation; A. Patel for technical assistance; J. K. Burks for discussions and technical assistance with image analysis; C. M. Johnston; H. Kimhi; S. J. Bronson; E. Kimhi and D. M. Aten for artistic work; and our patients and their families. This work was supported by National Institute of Dental and Craniofacial Research grant 5R01 DE014613 12 (J.N.M.) and by S.I.A. funds (G.A.C.). Y.C. and R.W. were funded by the National Natural Science Foundation of China (NSFC, 81741082). D.A.S. is supported by NCI fellowship NIH/NCI F30CA228258.

Author contributions J.W. and X.R. analysed the RNA and miRNA sequencing and performed statistical analysis; M.S. performed the ncRNA array experiments; C.I. analysed the RNA and miRNA arrays and performed statistical analysis; F.O.G.-N. analysed the TCGA data and performed statistical analysis; M.Z. produced and cultured oral keratinocytes; M.A., H.T., M.P.D., A.L. and S.A. performed in vivo orthotopic model experiments; M.A. and H.T. performed the animal surgery; Y.C. and R.W. performed in vivo carcinogen-induced genetically engineered mouse model experiments; H.T., M.P.D., A.L. and S.A. performed western blots; A.E.-N. and M.A. reviewed human and mouse pathology; P.M.D. provided and cultured human DRG neurons; M.A., H.T. and A.L. performed in vitro neuron growth studies and catecholamine measurements; A.L. performed miRNA qPCR; E.K. and M.P.D. designed and prepared CRISPR knockout cells; S.A. performed EV characterization and quantification; A.A.O. provided p53-isogenic cells; C.C. provided transgenic animals; D.A.S. interpreted p53 mutational data, analysed miRNA target gene pathways, and revised the manuscript; S.T. and C.R.P. provided the cohort of patients; A.Z. and E.T. provided miRNA-loaded nanoparticles (liposomes); M.A. wrote the manuscript with input from all authors; H.T. designed the figures; and G.A.C. and J.N.M. designed and supervised all experiments, prepared figures, and wrote the manuscript.

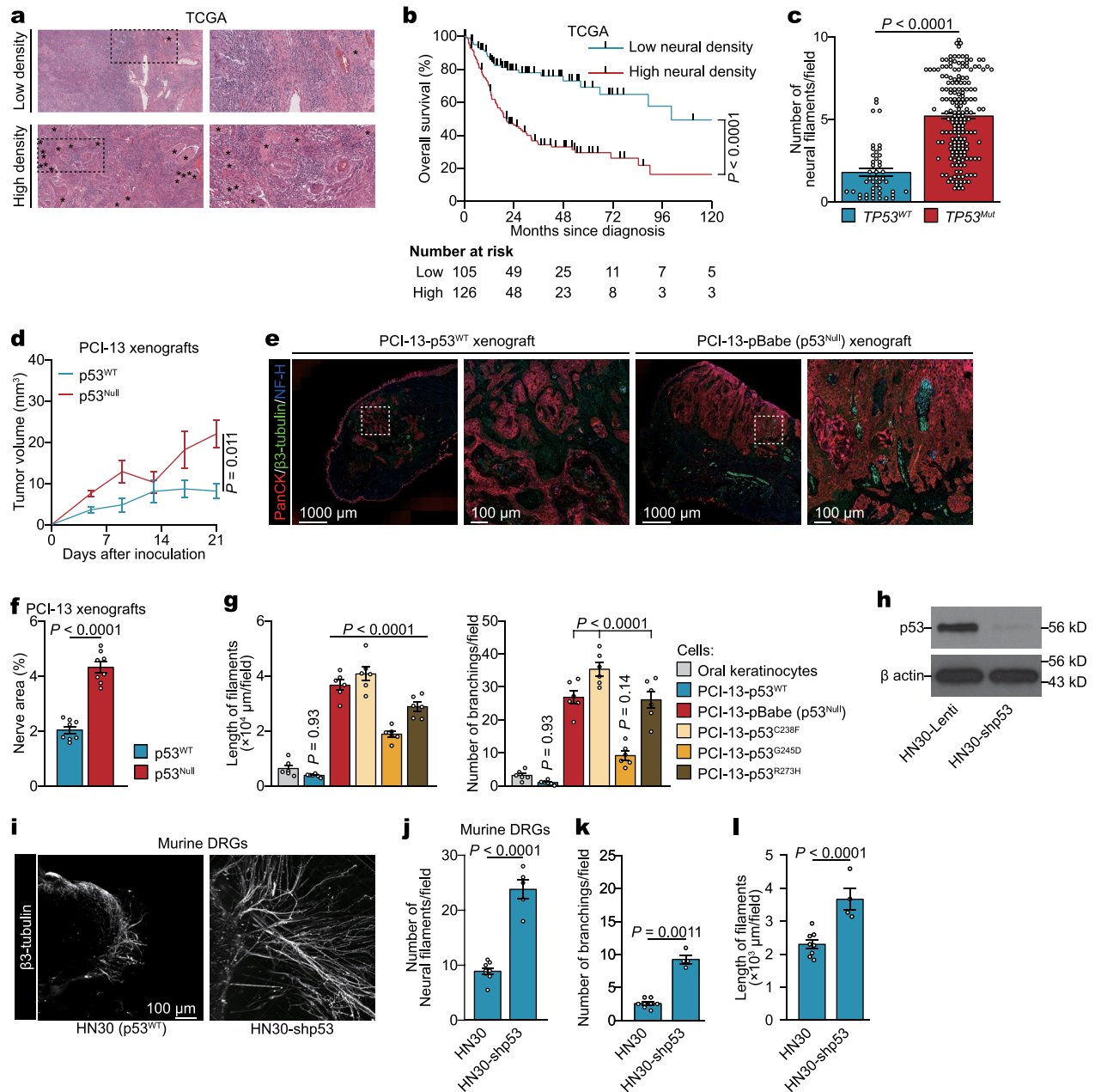
Competing interests The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-1996-3>.

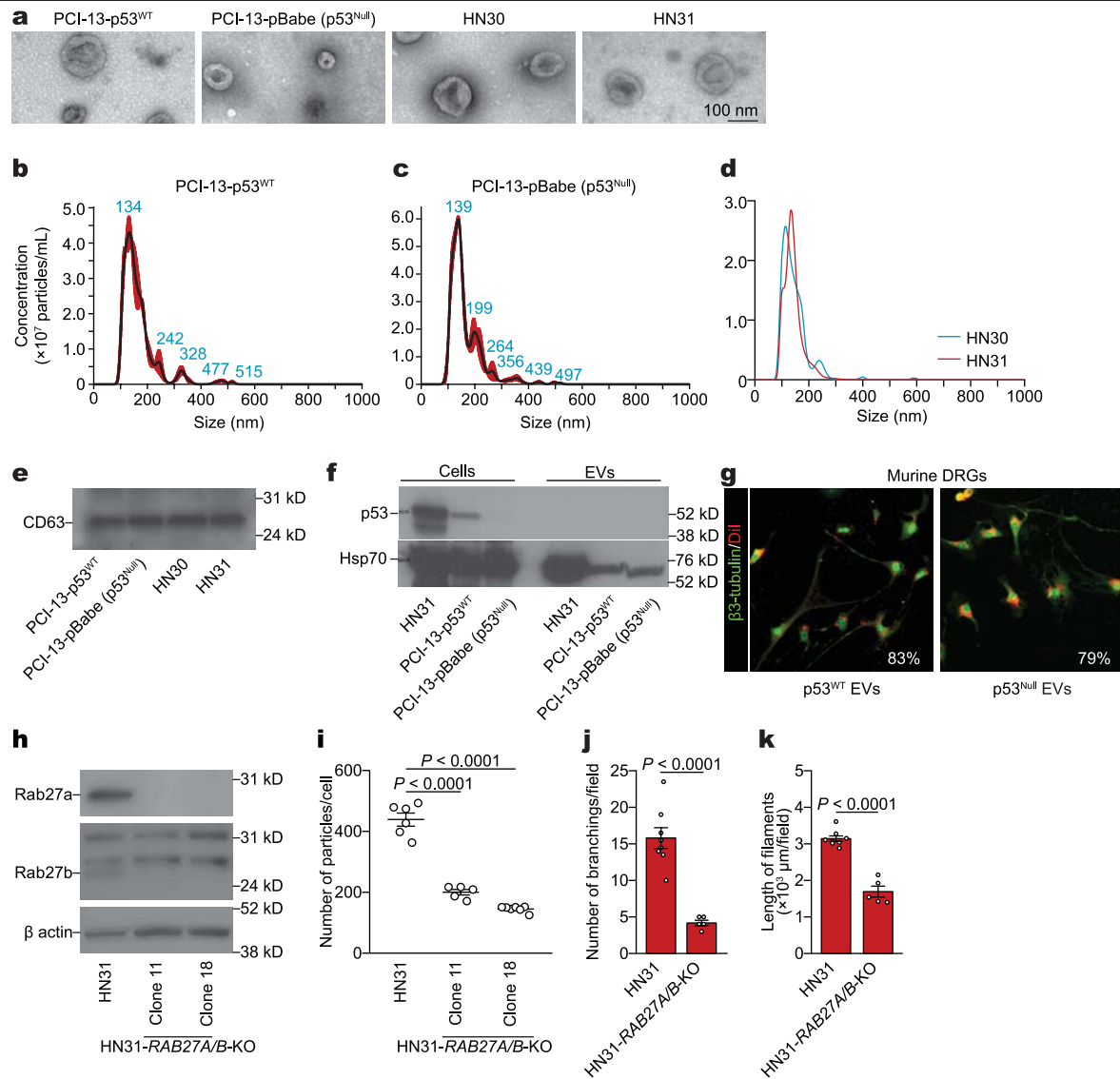
Correspondence and requests for materials should be addressed to M.A., G.A.C. or J.N.M.
Peer review information Nature thanks Elsa Flores, Hector Peinado and Hongjun Song for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | High nerve density was associated with the presence of $p53$ mutation. **a**, Representative haematoxylin and eosin image of OSCC samples from TCGA demonstrating low (top) and high (bottom) nerve densities; data independently replicated in 231 samples. Asterisks represent neural structures. **b**, Overall survival of patients with OSCC with high (≥ 4 neurofilaments per field) and low nerve densities. Two-sided log-rank test. **c**, Quantification of nerve density in TCGA OSCC patient cohort ($n = 231$). Bar graphs represent mean \pm s.e.m. Unpaired two-tailed t -test. **d**, Serial in vivo analyses of PCI-13 cell engraftment and growth in BALB/c (nu/nu) mice ($n = 6$ per group). Tumour growth curves represent mean tumour volume \pm s.e.m. Unpaired two-tailed t -test. **e**, Representative immunofluorescence of glossectomy specimens taken 4 weeks after orthotopic injection of isogenic $p53^{WT}$ or $p53^{Null}$ PCI-13 cells; data independently replicated in 16 mice.

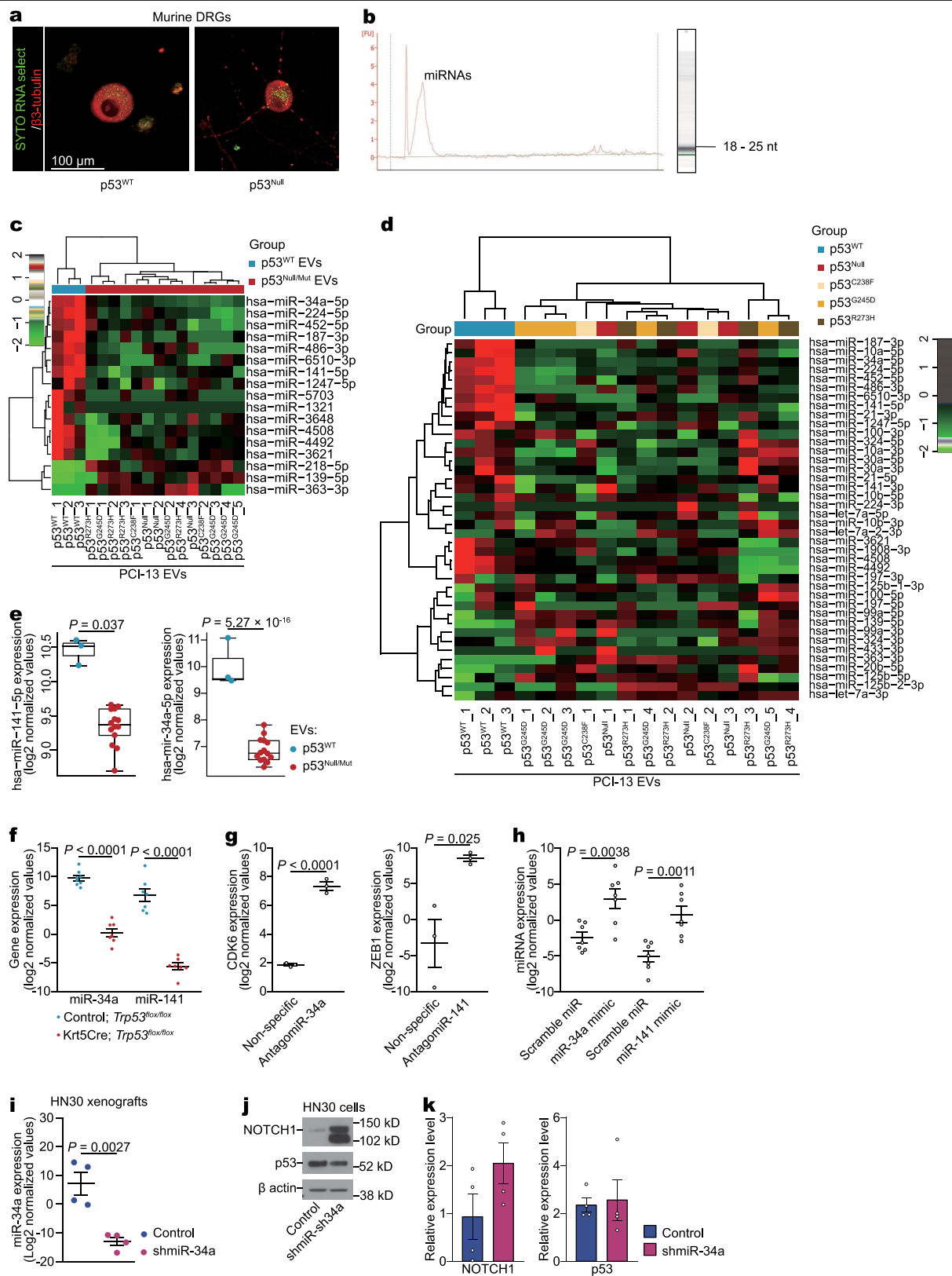
f, Quantification of nerve area in $p53^{WT}$ and $p53^{Null}$ OSCC xenografts ($n = 8$ mice per group). **g**, Quantification of neuritogenesis in DRG co-cultured with $p53$ -isogenic PCI-13 cells or normal oral keratinocytes ($n = 6$ biologically independent ganglia per cell line). **h**, Immunoblots demonstrating the knockdown of $TP53$ in human HN30 OSCC cells; data replicated in two independent experiments. **i**, Representative immunofluorescence staining of neo-neurites (β 3-tubulin⁺) in DRG co-cultured with HN30 (left) and HN30-shp53 (right) OSCC cells. Data independently replicated in 13 samples. **j–l**, In vitro quantification of number (**j**), branching (**k**), and length (**l**) of neurofilaments protruding from ex vivo DRG co-cultured with HN30 ($n = 8$ ganglia) or HN30-shp53 ($n = 5$ ganglia) OSCC cells. Bar graphs represent mean \pm s.e.m. One-way ANOVA with Tukey multiple comparison.



Extended Data Fig. 2 | OCSCC-derived EVs and neuritogenesis.

a, Representative transmission electron microscopy image of EVs from isogenic PCI-13, HN30, and HN31 OCSCC cells; data replicated in two independent experiments. **b–d**, Size distribution from nanoparticle tracking analysis of particles derived from p53^{WT} (**b**) or p53^{Null} (**c**) PCI-13 cells or HN30 and HN31 (**d**) cells; data replicated in two independent experiments. **e**, Western blot of the EV marker CD63 in EVs from PCI-13 and HN30/HN31 cells; data replicated in two independent experiments. **f**, Western blot of the p53 and controls (HSP70) in p53^{WT} or p53^{Null} PCI-13 and HN31 cells and their corresponding cell-derived EVs; data replicated in three independent experiments. **g**, Confocal immunofluorescence images showing EVs (lipophilic Dil-labelled, red) in the cytoplasm of a neuron (labelled with β 3-tubulin, green) 8 h after application of EVs derived from p53^{WT} or p53^{Null} PCI-13 cells. Percentage represents the proportion of Dil⁺ β 3-tubulin⁺ neurons out of all β 3-tubulin⁺ neurons ($n = 6$

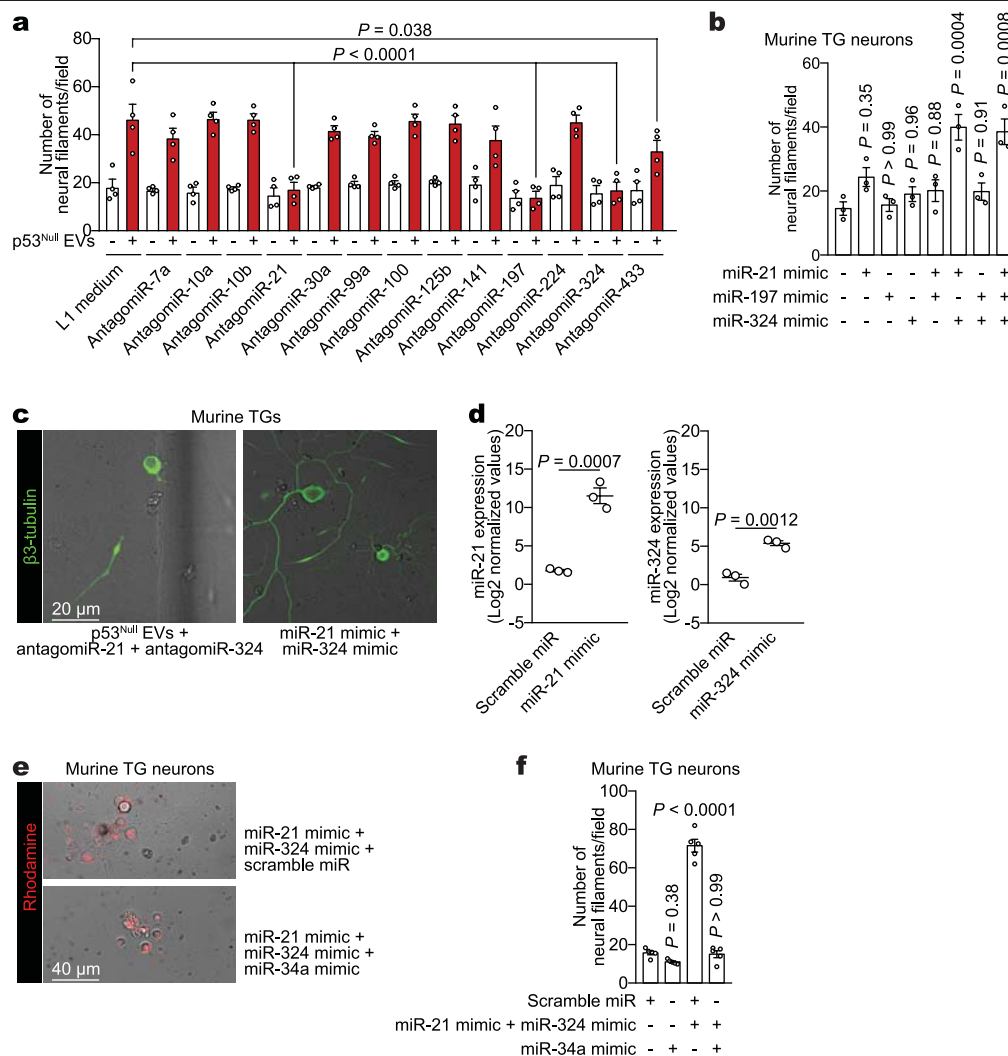
ganglia per condition). **h**, Immunoblots demonstrating the knockout (KO) of RAB27A and RAB27B in OCSCC cells edited with sgRNAs targeting *RAB27A* and *RAB27B* (HN31 clones 11 and 18, respectively) compared with HN31 controls; data replicated in two independent experiments. **i**, Nanoparticle tracking analysis of EV particle number in conditioned medium from HN31 clones 11 ($n = 5$ biologically independent samples) and 18 ($n = 7$ biologically independent samples) compared with HN31 controls ($n = 6$ biologically independent samples). Number of EVs was adjusted to cell number; bars represent mean \pm s.e.m. Unpaired two-tailed *t*-test. **j**, **k**, In vitro quantification of branching (**j**) and neurofilament length (**k**) in freshly collected DRGs cultured with conditioned medium from HN31 *RAB27A*^{+/+}*RAB27B*^{+/+} ($n = 8$) and HN31 *RAB27A*^{-/-}*RAB27B*^{-/-} ($n = 5$) isogenic human OCSCC cells. Bar graphs and tumour growth curves represent mean tumour volume \pm s.e.m. Unpaired two-tailed *t*-test.



Extended Data Fig. 3 | See next page for caption.

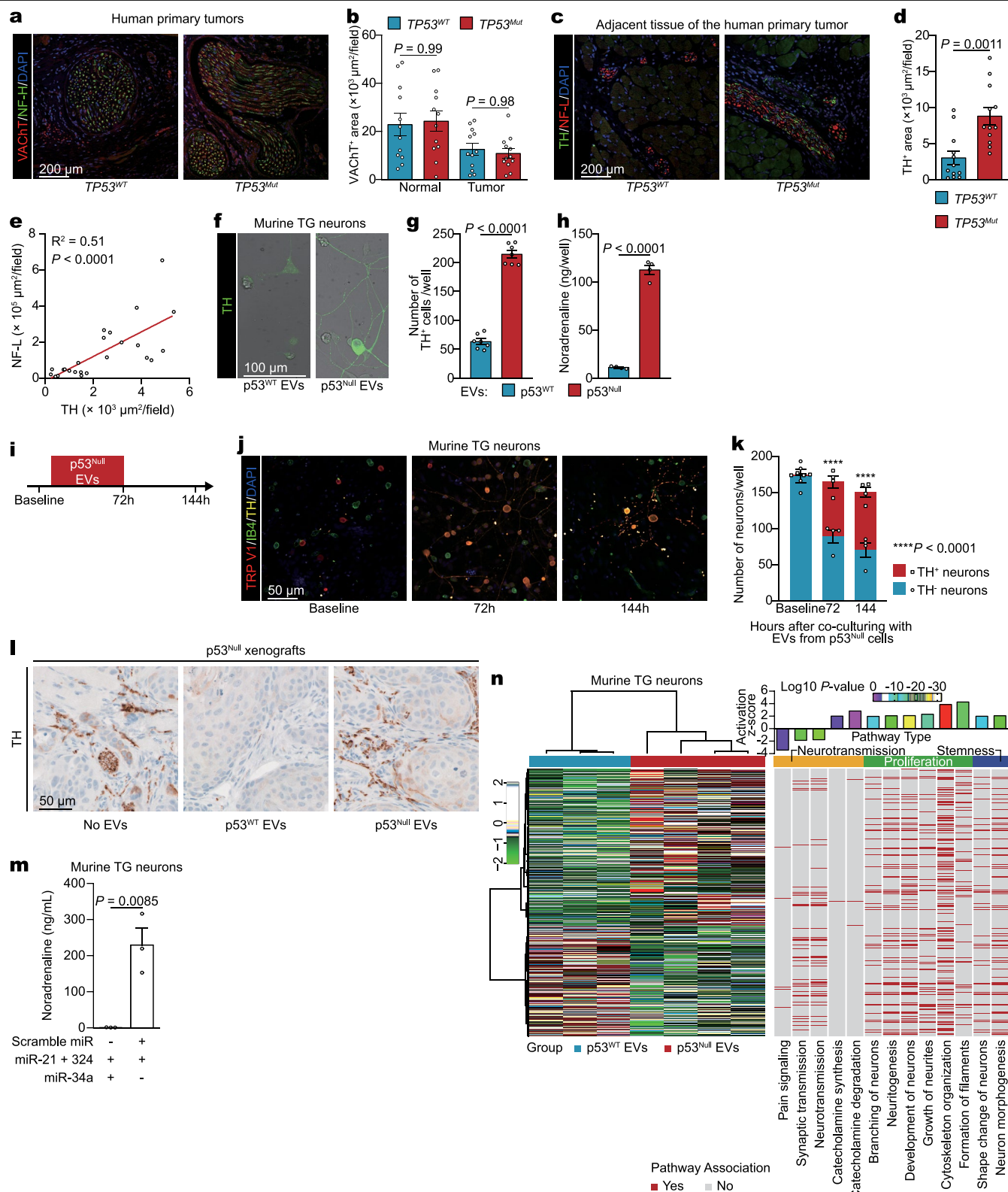
Extended Data Fig. 3 | p53-dependent miRNA in OSCC. **a**, OSCC RNA transfer to neurons via EVs. Representative confocal immunofluorescence image demonstrating PCI-13 cell-derived RNA labelled with SYTO RNASelect (green) in the perinuclear cytoplasm of a neuron (labelled with β 3-tubulin, red). Images were captured 12 h after application of EVs derived from PCI-13 cells labelled with SYTO RNASelect; data replicated in two independent experiments. **b**, EVs derived from PCI-13 cells contained mainly small RNA species. Bioanalyzer results showing presence of RNA in EVs from PCI-13 cells. Representative band of EV RNA by Agilent RNA Pico Chips; data independently replicated in ten experiments. **c**, An unsupervised hierarchical clustering heat map showing differentially expressed EV miRNAs between p53-isogenic PCI-13 cells. p53^{WT}, $n = 3$ biologically independent samples; p53^{null} and p53^{mut}, $n = 14$ biologically independent samples. **d**, Heat map of differentially expressed miRNA, arranged by unsupervised hierarchical clustering, presenting the miRNA sequencing for EVs derived from isogenic PCI-13 cells expressing p53^{WT} versus no p53 (p53^{null}) or mutant p53 (p53^{C238F}, p53^{G245D}, and p53^{R273H}). The Pearson distance and Ward's minimum variance method were used for pairwise clustering (**c**, **d**). Red and green indicate increased and decreased expression levels, respectively ($n = 2$ to 5 per group). **e**, Fold change in hsa-miR-141-5p and hsa-miR-34a-5p in EVs derived from p53^{WT} PCI-13 cells (blue, $n = 3$ biologically

independent samples) compared with p53^{null} or p53^{mut} cells (red, $n = 14$ biologically independent samples). Results are log₂ normalized. **f**, Real-time PCR quantification of miR-34a and miR-141 in ventral tongues from *Trp53^{fllox/fllox}* and *Krt5^{Cre}Trp53^{fllox/fllox}* mice ($n = 7$ per group). **g**, Real-time PCR quantification of CDK6 (miR-34a target) and ZEB1 (miR-141 target) in neurons treated with antagomiR-34 or antagomiR-141 compared with nonspecific antagomiR-treated controls ($n = 3$ biologically independent samples per group). **h**, Quantitative validation of miR-34a and miR-141 overexpression after transfection with miR-34a and miR-141 mimics, respectively. TG neurons were transfected with miR-34a mimic, miR-141 mimic, or scramble miR, and overexpression of miR-34a and miR-141 was confirmed by real-time PCR ($n = 7$ biologically independent samples per group). **i**, Real-time PCR quantification of miR-34a in orthotopic tumour xenografts of HN30 OSCC cells treated with shControl (blue) or shmiR-34a (purple). $n = 4$ biologically independent samples per group. **j**, **k**, Western blot of NOTCH1 (confirmed miR-34a target) in OSCC transfected with lentiviral miR-34a inhibitor or scramble miRNA inhibitor (**j**). Bar graph quantification of the blots demonstrates no impact of miR-34a inhibition on p53 expression and is normalized to the total amount of β -actin ($n = 4$ biologically independent samples per group, **j**). Unpaired two-tailed t -test; bars and dot plots represent mean \pm s.e.m. (**e**–**i**, **k**).



Extended Data Fig. 4 | microRNAs modulate neuritogenesis. **a**, Screening of candidate neuritogenesis-associated miRNAs. Quantification of neuritogenesis 72 h after neuron–EV co-culture. Eight hours after transfection with 13 different antagonomiRs, TG neurons were incubated with EVs derived from p53^{null} PCI-13 cells ($n = 4$ biologically independent samples per condition). One-way ANOVA with Tukey multiple comparison. **b**, Quantification of neuritogenesis in TG neurons 72 h after transfection with miR-21 mimic, miR-197 mimic, or miR-324 mimic or co-transfection with their combinations ($n = 3$ biologically independent samples per condition). **c**, Representative fluorescent–bright-field overlay images demonstrating the lack of response of TG neurons exposed to EVs derived from p53^{null} PCI-13 cells after co-transfection with antagonomiR-21 and antagonomiR-324 and the response of TG

neurons after miR-21 and miR-324 mimic co-transfection. Data replicated across six independent samples. **d**, Quantitative validation of miR-21 and miR-324 overexpression in TG neurons incubated with liposomes containing miR-21, miR-324, and scramble miRNA ($n = 3$ biologically independent samples per group). **e**, Representative fluorescence–bright-field overlay images of TG neurons exposed to liposomes containing miR-21, miR-324, and scramble miRNA or liposomes containing miR-21, miR-324, and miR-34a; data independently replicated in 20 wells. **f**, Quantification of neuritogenesis in TG neurons 72 h after neuron–liposome co-culture ($n = 5$ biologically independent samples per condition). Unpaired two-tailed *t*-test; bars represent mean \pm s.e.m. (**a**, **b**, **d**) or one-way ANOVA with Tukey multiple comparisons (**f**).

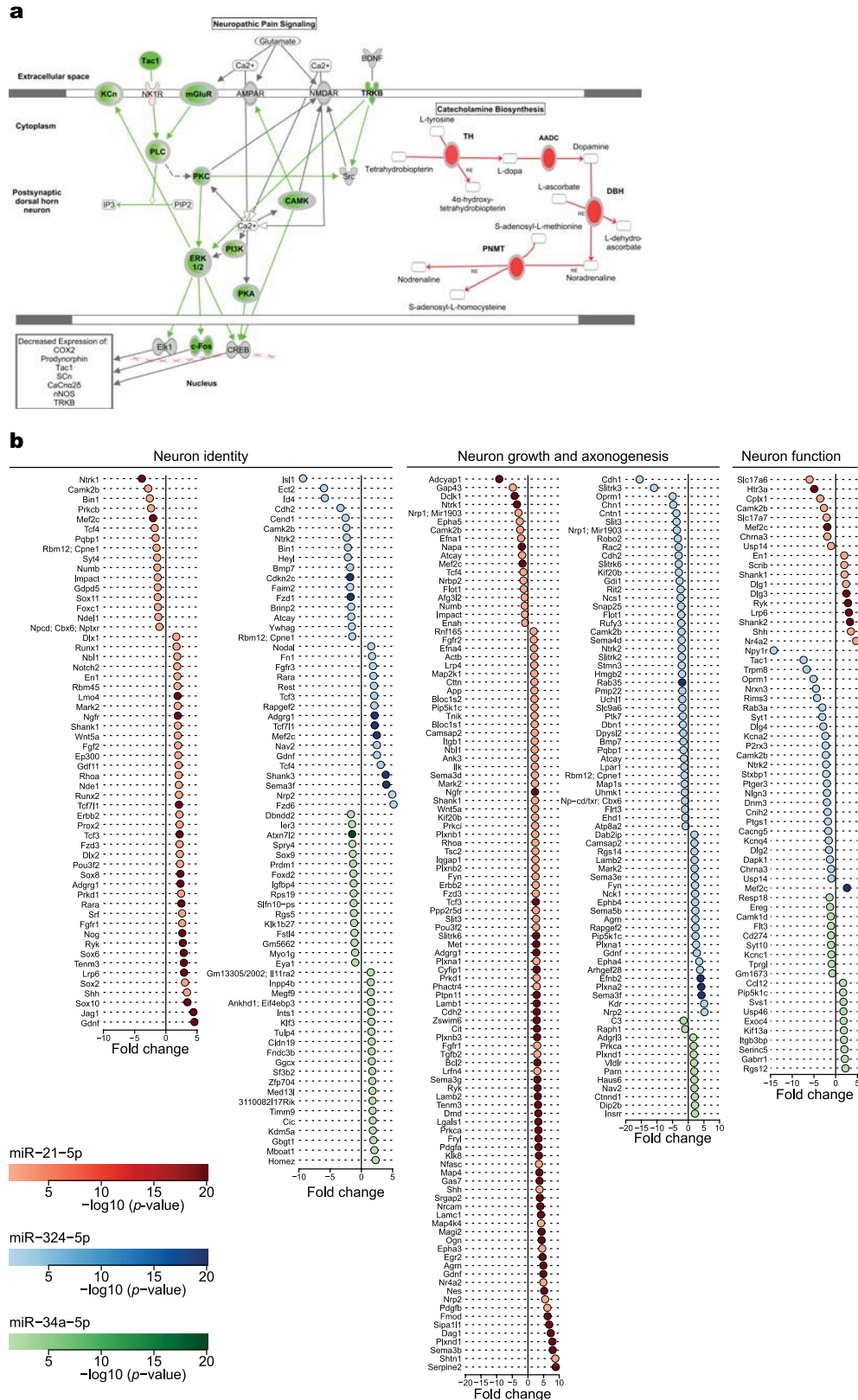


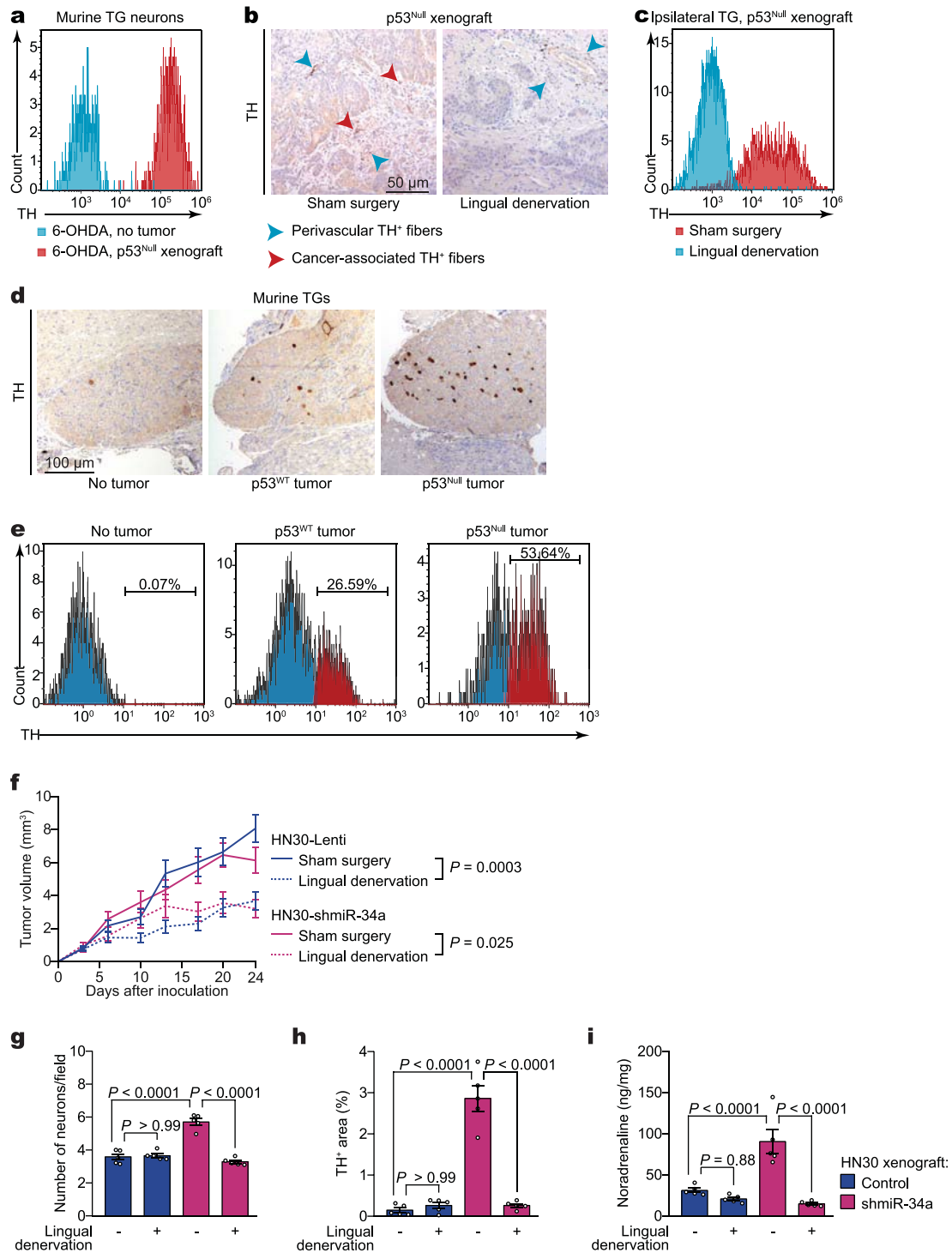
Extended Data Fig. 5 | See next page for caption.

Article

Extended Data Fig. 5 | *TP53* deficiency does not change parasympathetic nerve fibre densities in human OCSCC specimens. **a, b**, Representation of vesicular acetylcholine transporter (VACHT)⁺ nerve densities in both *TP53*^{WT} and *TP53*^{mut} OCSCC tissues; data independently replicated in 24 patient specimens (**a**). Quantification of cholinergic VACHT⁺ neural areas in *TP53*-sufficient (*TP53*^{WT}, blue, *n* = 12) and *TP53*-deficient (*TP53*^{mut}, red, *n* = 12) human OCSCC tissues. Each dot represents the mean for one patient (**b**). NF-H, neurofilament heavy. **c, d**, *TP53* deficiency increases the sympathetic nerve fibre density in normal tongue tissue surrounding OCSCC in humans. Representative images showing TH⁺ adrenergic neural fibres in human normal tongue tissue surrounding OCSCC with *TP53*^{WT} (left) or *TP53*^{mut} (right) (TH, green; neurofilament light (NF-L), red; DAPI, blue). Data independently replicated in 24 patient specimens (**c**). Quantification of adrenergic TH⁺ areas in *TP53*-sufficient (*TP53*^{WT}, blue, *n* = 12) and *TP53*-deficient (*TP53*^{mut}, red, *n* = 12) human OCSCC samples. Each dot represents the mean for a patient (**d**). **e**, Correlation of TH and NF-L expression levels. Linear regression (r^2 , *n* = 24 biologically independent samples). **f–h**, Representative images of TG neurons labelled with anti-TH antibody after incubation with EVs derived from p53^{WT} or p53^{null} PCI-13 cells; data independently replicated in 14 wells (**f**). Quantification

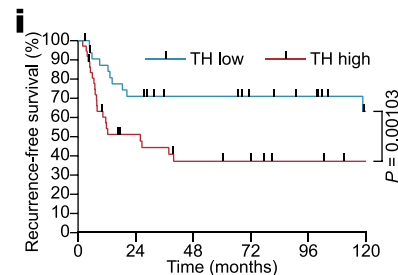
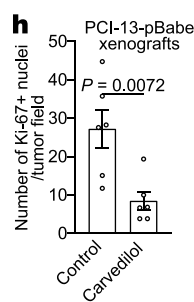
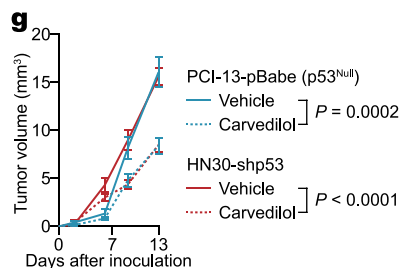
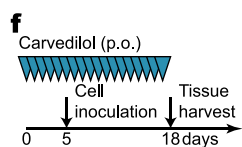
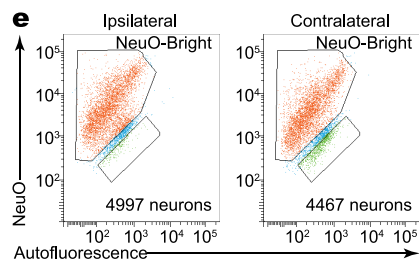
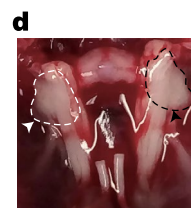
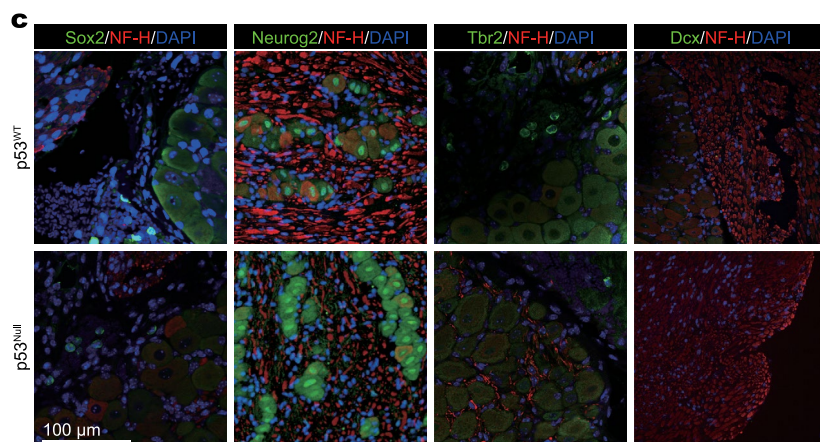
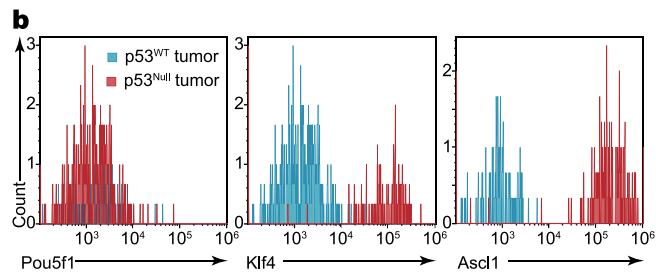
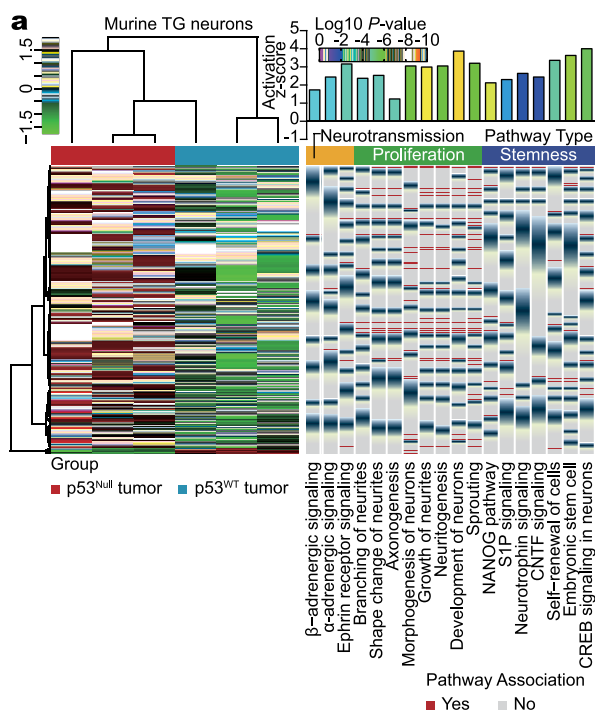
of TH⁺ neurons (*n* = 7 biologically independent samples per condition, **g**), and noradrenaline levels (*n* = 4 biologically independent samples per condition, **h**). **i–k**, Co-culture of TG neurons with p53^{null} EVs for 72 h induced TH coexpression in TRPV1⁺ but not IB4⁺ neurons; TH expression remained stable 72 h after washout of the EVs. *n* = 4 biologically independent samples per condition (**k**). **l**, TH⁺ neural areas in PCI-13–p53^{null} orthotopic tumours injected daily with no EVs or with EVs derived from p53^{null} or p53^{WT} PCI-13 cells for 3 weeks; data independently replicated in 15 mice. **m**, Co-culture of TG with liposomes containing miR-21 and miR-324 but not miR-34a increases catecholamine synthesis. Noradrenaline levels in neurons cultured with nano-liposomes containing miR-21 + miR-34a + miR-324 or miR-21 + miR-324 controls, quantified by enzyme-linked immunosorbent assay. *n* = 3 biologically independent samples per condition. **n**, Heat map of differentially expressed genes in mouse TG neurons co-cultured with p53-isogenic EVs. Enriched Gene Ontology terms of the neurons were plotted at fold enrichment with the associated log *P* value (Fisher's exact algorithm for functional gene set enrichment); *n* = 3 biologically independent samples for p53^{WT} and *n* = 4 biologically independent samples for p53^{null}. Mean ± s.e.m.; unpaired two-tailed *t*-test (**b, d, g, h, k, m**).





Extended Data Fig. 7 | Loss of p53 in OCSCC induces adrenergic switch proximally in TG neurons. **a**, Flow cytometry quantification of neurotrophin-3-positive (NT3⁺), TH⁺ neurons in freshly collected ipsilateral TG neurons 3 weeks after orthotopic inoculation of p53^{Null} PCI-13 cells to the tongues of syngeneic mice. Non-tumour-bearing, syngeneic mice were used as controls (n = 6). **b**, Representative immunohistochemical analysis for TH⁺ in orthotopic xenografts; data independently replicated in 16 mice. **c**, Flow cytometry quantification of NT3⁺TH⁺ neurons in ipsilateral TG neurons (n = 6 mice per condition). **d**, Representative images of TH⁺ TG neurons in mice without tumours (left) and 3 weeks after injection of p53^{WT} (middle) or p53^{Null} (right) PCI-13 cells to the ipsilateral tongue; data independently replicated in nine mice. **e**, Flow cytometry quantification of NT3⁺ TH⁺ neurons in freshly

collected ipsilateral TG 3 weeks after orthotopic inoculation of p53^{WT} (middle) and p53^{Null} (right) PCI-13 cells to the tongue. Non-tumour-bearing mice were used as controls (left, n = 12 per group). **f**, Serial in vivo analyses of tumour growth after engraftment of HN30 transfected with either control lentivirus (HN30-lenti) or shmiR34a (HN30-shmiR34a) into BALB/c (nu/nu) mice. Mice were randomized and underwent lingual denervation or sham surgery 1 week before cell injection (n = 8 per group). Tumour growth curves represent mean tumour volume ± s.e.m.; unpaired two-tailed t-test. **g**–**i**, Neural density (**g**), TH⁺ area (**h**), and noradrenaline levels in vivo (**i**) in HN30-lenti and HN30-shmiR34a orthotopic xenografts with and without lingual denervation (n = 5 biologically independent samples per condition). Bars indicate mean ± s.e.m.; unpaired two-tailed t-test.



Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Characterization of OCSCC-induced neural transcriptional program. **a**, Heat map of differentially expressed genes arranged by unsupervised hierarchical clustering in TG neurons 3 weeks after orthotopic injection of p53^{null} or p53^{WT} PCI-13 cells ($n = 3$ biologically independent samples per condition) and enriched Gene Ontology terms plotted by fold enrichment with the associated log P value (right; Fisher's exact algorithm for functional gene set enrichment). **b**, Flow cytometry quantification of NeuroFluor-positive (NeuO⁺), POU5F1⁺ (left), NeuO⁺KLF4⁺ (middle), and NeuO⁺ASCL1⁺ (right) neurons in ipsilateral TGs after orthotopic injection of p53^{null} or p53^{WT} PCI-13 cells; data independently replicated in six mice. **c**, Representative images in freshly collected TG neurons (red, NF-H⁺) from BALB/c (*nu/nu*) mice after orthotopic injection of either p53^{null} or p53^{WT} PCI-13 cells to the ipsilateral tongue, demonstrating lack of nuclear expression of the transcription factors SOX2, TBR2, and DCX and similar neurogenin2 expression between groups. Data independently replicated in six mice; DAPI, blue. **d**, **e**, Representative necropsy photograph (TG delineated by dashed line

in **c**) and flow cytometry quantification of NeuO⁺ neurons in freshly collected TG 3 weeks after tumour injection to the left side of the tongue; ipsilateral (**c**, right, black arrowhead) and contralateral (**c**, left, white arrowhead) ganglia were similar in size. None of the tumours crossed the midline of the tongue ($n = 6$). **f**, **g**, Mice were treated daily with either β -adrenergic receptor blocker carvedilol or vehicle via oral gavage. On day 5, mice were orthotopically xenografted with human p53-deficient (p53^{null} PCI-13 or HN30-shp53) cells to the tongue. Serial in vivo tumour volume measurement ($n = 12$ per group except for p53^{null} PCI-13 tumour-bearing mice with carvedilol treatment, $n = 13$; **f**). **h**, Adrenergic inhibition decreases OCSCC proliferation in vivo. Carvedilol injections inhibited the proliferation of p53^{null} PCI-13 cells orthotopically implanted into the tongue, as determined by Ki-67 expression ($n = 6$ biologically independent samples per condition). **i**, Kaplan–Meier curves showing the recurrence-free survival of patients with high ($>2,000 \mu\text{m}^2$ per field) and low ($\leq 2,000 \mu\text{m}^2$ per field) TH⁺ adrenergic nerve densities. Mean \pm s.e.m.; unpaired two-tailed t -test.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- ☐ ☒ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☐ ☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☒ ☐ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data was collected using Microsoft Excel 2016 64-bit (Microsoft), Vevo 2100 ultrasound imaging system, iQ3 (Andor), the PerkinElmer Vectra platform, Axio Scan.Z1 (Zeiss), FACS Diva Version 8.0.1 (BD Biosciences), NanoSight LM10 Nanoparticle Characterization system, AMT Imaging System (Advanced Microscopy Techniques Corp.).

Data analysis

Data was analyzed using GraphPad Prism Version 7.03, SAS JMP Pro software version 12.1.0, the R language environment for statistical computing version 3.1.3, imageJ 1.52 (NIH, USA), the FilamentTracer module in the Imaris software (Bitplane), the Simple Neurite Tracer plugin in the Fiji build of ImageJ (NIH), TissueFinder in PerkinElmer inForm software, Pannoramic Viewer (3DHISTECH), FACSDiva 8.0 software (BD Biosciences), FlowJo 1.0.1, Kaluza software (Beckman Coulter), the package edgeR, QIAGEN Ingenuity Pathway Analysis. Western blots were cropped in Adobe Photoshop CC 2017 and Adobe Illustrator CC 2017. miRNA analysis: Cutadapt version 1.8.1 for trimming adapters from reads, BWA version 0.7.15 for mapping the short reads, SAMtools version 1.2 for manipulating bam/sam files featureCounts from Subread version 1.5.2 for miRNA quantification. mRNA analysis: murine: FastQC version 0.11.7 for read quality check, TopHat2 version 2.0.14 for mapping, SAMtools version 1.2 for manipulating bam/sam files, HTSeq version 0.6.1 for gene quantification; human: FastQC version 0.11.8 for read quality check, TopHat2 version 2.1.1 for mapping, SAMtools version 1.2 for manipulating bam/sam files, HTSeq version 0.11.0 for gene quantification.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Source data for the figures and extended data figures are provided. The datasets generated during the current study are available from the corresponding authors upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	<p>Samples sizes from other experiments were estimated from similar experiments in former publications of the group.</p> <p>Sample sizes for clinical data: No statistical methods were used to predetermine sample sizes and samples were selected based upon the availability of data as outlined below.</p> <p>-TCGA cohorts of head and neck cancer (oral cancer) were assessed.</p> <p>-TP53 status and the expression levels of tyrosine hydroxylase, vesicular choline transporter, neurofilament-light, and neurofilament-heavy in glossectomy tissues from treatment-naïve patients with tongue cancer treated at the University of Texas MD Anderson Cancer Center, Houston, Texas were analyzed (n = 24, 12 with wild-type TP53 and 12 with mutant TP53).</p> <p>-The expression levels of tyrosine hydroxylase within tumor areas in patients with oral cavity squamous cell carcinoma was evaluated and compared with their clinical characteristics and survival (n = 70).</p> <p>For human DRGs, experiments were prioritized based on the availability of the specimen.</p> <p>In vitro studies: sample sizes were determined based on the results of pilot studies, and previous similar studies that have given statistically significant results.</p> <p>In vivo studies: Animal experiments were conducted using between three and 13 mice per group, based on the results of pilot studies and previous studies (that have given statistical significant results based on the variance of xenograft growth in control mice. These power calculations indicated use of at least 3 mice per genotype to give 80% power to detect an effect size of 20% with a significance level of 0.05).</p>
Data exclusions	No data excluded.
Replication	<p>All attempts at replication were successful. Experiments were performed at least two times and/or with sufficient cells/animals per group to demonstrate statistical significance. Number of replicates of each experiment is indicated in the corresponding figure legend.</p> <p>Data was replicated using:</p> <ol style="list-style-type: none"> 1. Human specimen (cancer and neurons) 2. Orthotopic xenograft models (HN30, HN31, PCI13, p53- and miR34- isogenic cell lines) 3. Transgenic (Trp53) model <p>We have designed (CRISPR/Cas9) 2 independent clones (#11 and #18) of RAB27A-/-;RAB27B-/- knock out cells.</p> <p>We have designed p53 function studies using genetic approach with 5 isogenic PCI-13 (exogenous p53) cell lines and 2 isogenic HN30 (endogenous p53) OCSCC cell lines.</p>
Randomization	Animals were randomly allocated for surgical denervation or carvedilol/6OHDA treatment prior to surgery or treatment initiation, respectively. For exosome injection studies and orthotopic models with no intervention (other than tumor engraftment) animals were randomly allocated prior to cell inoculation.
Blinding	The researchers were blinded during the measurement of the tumor size and body weight of mice except in the experiment shown in Fig. 3i-k. The researchers were blinded during outcome assessment.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used

Rat anti-human/mouse Oct3/4-APC (Clone:240408, R&D systems, cat#:IC1759A, 1:400).
 Mouse anti-human CD63 (unconjugated, Clone TS63, Abcam, cat#:ab59479, 1:1000, Lot:GR3186539-1).
 Rabbit anti-TRPV1 (unconjugated, Clone VR1, Alomone Labs cat#:ACC-030; 1:500, Lot:ACC030AG1040).
 Rabbit anti-doublecortin (unconjugated, polyclonal, Abcam, cat#:ab18723, 1:1000, Lot#: GR3224908-1).
 Rabbit anti-Rab27B (unconjugated, polyclonal, Abcam, cat#:ab103418, 1:200, Lot #: GR3238038-3).
 Mouse anti-Rab27A (unconjugated, monoclonal, Abcam, cat#:ab55667, 1:200, Lot #: GR3198959-1).
 Rabbit anti-TH (unconjugated, polyclonal, Bioss, cat#:bs-0016R-A488, 1:50, Lot:AE092200).
 Mouse anti-TP53 (unconjugated, Clone DO-1, Cell Signaling, cat#:18032, 1:1000).
 Rabbit anti-Notch1 (unconjugated, Clone D1E11, Cell Signaling, cat#:3608, 1:1000).
 Mouse anti- β actin (unconjugated, Clone AC-15, Sigma, cat#:A1978, 1:4000).
 Mouse anti neuron-specific β 3-tubulin-APC (Clone TUJ-1, R&D systems, cat#:IC1195A, 1:100, Lot: ABFK0117091).
 Rabbit anti-Oct4 (unconjugated, polyclonal, Abcam, cat#:ab18976, 1:100, Lot:GR3202710-3).
 Rabbit anti- β 3-tubulin (polyclonal, unconjugated, Abcam, cat#:ab18207, 1:2000, Lot#: GR3221401-3 and GR3196636-1).
 Chicken anti-neurofilament heavy (polyclonal, unconjugated, Abcam, cat#:ab4680, 1:1000, Lot:GR3241438-4).
 Rabbit anti-MASH1/ASCL1 (polyclonal, unconjugated, Abcam, cat#:ab74065; 1:250, Lot:GR3178505-1).
 Rabbit anti-TH (polyclonal, unconjugated, EMD Millipore, cat#: AB152, 1:400, Lot:2971004, 3031639 and 3072361).
 Rabbit anti-neurogenin2 (unconjugated, polyclonal, EMD Millipore, cat#:AB5682, 1:400, Lot:3066197).
 Rabbit anti-neurofilament heavy polypeptide (unconjugated, polyclonal, Abcam, cat#:ab8135, 1:1000, Lot:GR3224043-9).
 Chicken anti-68kDa NF/NF-L (unconjugated, polyclonal, Abcam, cat#:ab24520, 1:700, Lot:GR3206616-5).
 Mouse anti-SOX2 (unconjugated, Clone 245610, R&D systems, cat#: IC2018V-100UG, 1:500, Lot:1502977).
 Rabbit anti-neurofilament L (unconjugated, polyclonal, EMD Millipore, cat#:AB9568, 1:700, Lot:2943223).
 Goat anti-VACHT (unconjugated, polyclonal, EMD Millipore, cat#:ABN100, 1:1000, Lot: 2899777).
 Chicken anti-Tbr2 (unconjugated, polyclonal, EMD Millipore, cat#:AB15894, 1:500, Lot:3071578).
 Mouse anti-pancytokeratin (unconjugated, Polyclonal PAN-CK, ThermoFisher, cat#:MA5-13203, 1:50).
 Rabbit anti-KLF4 (unconjugated, polyclonal Novus, cat#:NBP2-24749, 1:100, Lot:05232589B-07).

Validation

Well-validated human and mouse antibodies were purchased from established commercial vendors, including Abcam, Sigma, EMD Millipore, Novus, Bioss, R&D Systems, Cell Signaling, and Life Technologies. Unless otherwise noted, antibodies were used at manufacturer- and primary literature-validated concentrations for the relevant assays, as detailed below:
 Rat anti-human/mouse Oct3/4-APC (Clone:240408, R&D systems, cat#:IC1759A, 1:400). Mutnal MB et al. Murine cytomegalovirus infection of neural stem cells alters neurogenesis in the developing brain. PLoS ONE, 6(1):e16211 (2011). Validated data in FC by the provider.
 Mouse anti-human CD63 (unconjugated, Clone TS63, Abcam, cat#:ab59479, 1:1000, Lot:GR3186539-1), Matsuzaki K et al. MiR-21-5p in urinary extracellular vesicles is a novel biomarker of urothelial carcinoma. Oncotarget 8:24668-24678 (2017). Zonneveld MI et al. Recovery of extracellular vesicles from human breast milk is influenced by sample collection and vesicle isolation procedures. J Extracell Vesicles 3:N/A (2014). Validated data in IHC, WB, IF by the provider.
 Rabbit anti-TRPV1 (unconjugated, Clone VR1, Alomone Labs cat#:ACC-030; 1:500, Lot:ACC030AG1040), Devesa et al. α CGRP is essential for algescic exocytotic mobilization of TRPV1 channels in peptidergic nociceptors. Proc Natl Acad Sci U S A. 111(51):18345-50 (2014). Knockout-validated by provider.
 Rabbit anti-doublecortin (unconjugated, polyclonal, Abcam, cat#:ab18723, 1:1000, Lot#: GR3224908-1), Kovalchuk Y et al. In vivo odourant response properties of migrating adult-born neurons in the mouse olfactory bulb. Nat Commun 6:6349 (2015). Validated data in IHC, WB, FC, IF by the provider.
 Rabbit anti-Rab27B (unconjugated, polyclonal, Abcam, cat#:ab103418, 1:200, Lot #: GR3238038-3). Jiang Y et al. MicroRNA-599 suppresses glioma progression by targeting RAB27B. Oncol Lett 16:1243-1252 (2018). Validated data in WB by the provider.
 Mouse anti-Rab27A (unconjugated, monoclonal, Abcam, cat#:ab55667, 1:200, Lot #: GR3198959-1). Uchino K et al. Therapeutic Effects of MicroRNA-582-5p and -3p on the Inhibition of Bladder Cancer Progression. Mol Ther 21:610-9 (2013). Knockout-validated data in WB by the provider.
 Rabbit anti-TH (unconjugated, polyclonal, Bioss, cat#:bs-0016R-A488, 1:50, Lot:AE092200). Validated data in IHC, WB, FC by the provider. Additional supportive validation exist in Antibodypedia.
 Mouse anti-TP53 (unconjugated, Clone DO-1, Cell Signaling, cat#:18032, 1:1000). Validated data in WB by the provider.
 Rabbit anti-Notch1 (unconjugated, Clone D1E11, Cell Signaling, cat#:3608, 1:1000). Man, Jianghong, et al. Hypoxic induction of vasorin regulates Notch1 turnover to maintain glioma stem-like cells. Cell stem cell. 22.1:104-118 (2018). Validated data in WB by the provider.
 Mouse anti- β actin (unconjugated, Clone AC-15, Sigma, cat#:A1978, 1:4000). Validated data in WB by the provider.
 Mouse anti neuron-specific β 3-tubulin-APC (Clone TUJ-1, R&D systems, cat#:IC1195A, 1:100, Lot: ABFK0117091). Noristani HN et

al. Spinal cord injury induces astroglial conversion towards neuronal lineage. *Mol Neurodegener.* 11(1):68 (2016). Validated data in FC by the provider.

Rabbit anti-Oct4 (unconjugated, polyclonal, Abcam, cat#:ab18976, 1:100, Lot:GR3202710-3), Hassiotou F et al. Expression of the Pluripotency Transcription Factor OCT4 in the Normal and Aberrant Mammary Gland. *Front Oncol* 3:79 (2013). Lee SJ et al. Adult stem cells from the hyaluronic acid-rich node and duct system differentiate into neuronal cells and repair brain injury. *Stem Cells Dev* 23:2831-40 (2014). Validated data in IHC, WB by the provider.

Rabbit anti- β -tubulin (polyclonal, unconjugated, Abcam, cat#:ab18207, 1:2000, Lot#: GR3221401-3 and GR3196636-1), Delgado-Esteban M et al. APC/C-Cdh1 coordinates neurogenesis and cortical size during development. *Nat Commun* 4:2879 (2013). Validated data in IHC, WB, FC, IF by the provider.

Chicken anti-neurofilament heavy (polyclonal, unconjugated, Abcam, cat#:ab4680, 1:1000, Lot:GR3241438-4). Wirt SE et al. G1 arrest and differentiation can occur independently of Rb family function. *J Cell Biol* 191:809-25 (2010). Validated data in IHC, WB, IF by the provider.

Rabbit anti-MASH1/ASCL1 (polyclonal, unconjugated, Abcam, cat#:ab74065; 1:250, Lot:GR3178505-1). Validated data in IHC, WB, IF by the provider.

Rabbit anti-TH (polyclonal, unconjugated, EMD Millipore, cat#: AB152, 1:400, Lot:2971004, 3031639 and 3072361), Magnon C et al. Autonomic Nerve Development Contributes to Prostate Cancer Progression. *Science* 341(6142):1236361 (2013). Validated data in IHC, ELISA WB, IF by the provider.

Rabbit anti-neurogenin2 (unconjugated, polyclonal, EMD Millipore, cat#:AB5682, 1:400, Lot:3066197), Validated data in IHC, ELISA WB, IF by the provider.

Rabbit anti-neurofilament heavy polypeptide (unconjugated, polyclonal, Abcam, cat#:ab8135, 1:1000, Lot:GR3224043-9); Woo SH et al. Piezo2 is required for Merkel-cell mechanotransduction. *Nature* 509:622-6 (2014). Validated data in IHC, WB, IF by the provider.

Chicken anti-68kDa NF/NF-L (unconjugated, polyclonal, Abcam, cat#:ab24520, 1:700, Lot:GR3206616-5), Validated data in IHC, WB, IF by the provider.

Mouse anti-SOX2 (unconjugated, Clone 245610, R&D systems, cat#: IC2018V-100UG, 1:500, Lot:1502977). Najm, FJ et al. Transcription factor-mediated reprogramming of fibroblasts to expandable, myelinogenic oligodendrocyte progenitor cells. *Nat Biotechnol*, 31(5):426-33 (2013). Validated data in WB, FC, IF by the provider.

Rabbit anti-neurofilament L (unconjugated, polyclonal, EMD Millipore, cat#:AB9568, 1:700, Lot:2943223), Validated data in IHC by the provider.

Goat anti-VACHT (unconjugated, polyclonal, EMD Millipore, cat#:ABN100, 1:1000, Lot: 2899777), Validated data in IHC by the provider.

Chicken anti-Tbr2 (unconjugated, polyclonal, EMD Millipore, cat#:AB15894, 1:500, Lot:3071578); He Y et al. ALK5-dependent TGF- β signaling is a major determinant of late-stage adult neurogenesis. *Nature neuroscience*, 17:943-52 (2014). Validated data in WB, IHC by the provider.

Mouse anti-pancytokeratin (unconjugated, Polyclonal PAN-CK, ThermoFisher, cat#:MA5-13203, 1:50), Validated data in WB, IHC, IF by the provider. Additional supportive validation exist in Antibodypedia.

Rabbit anti-KLF4 (unconjugated, polyclonal Novus, cat#:NBP2-24749, 1:100, Lot:05232589B-07). Validated data in WB, IHC, FC by the provider. Additional supportive validation exist in Antibodypedia.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	PCI-13 cells were obtained from the laboratory of Dr. Jennifer Grandis (University of Pittsburgh, Pittsburgh, PA). HN30 and HN31 cells were obtained from the laboratory of Dr. John Ensley (Wayne State University, Detroit, MI) and from Dr. Barbara Frederick (University of Colorado Health Sciences Center-Colorado), respectively.
Authentication	All human cell lines were authenticated upon arrival by STR profiling using 14 short tandem repeat (STR) loci including the gender determining locus, Amelogenin.
Mycoplasma contamination	All cell lines were tested for mycoplasma contamination periodically, including immediately upon receipt via the MycoAlert Mycoplasma Testing Kit (Lonza). Results were always negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in this study.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	Male or female B6.129P2-Trp53tm1Brn/J, and BALB/c nu/nu (B6.Cg-Foxn1nu+/-) mice at the age of 6 to 8 weeks were obtained from the Jackson Laboratory. Krt5-Cre was obtained from Dr. Carlos Caulin.
Wild animals	No wild animals were used.
Field-collected samples	No field-collected samples were used.
Ethics oversight	Study protocols were approved by the University of Texas MD Anderson Cancer Center Institutional Animal Care and Use Committee (IACUC).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics

Historical specimen of patients who underwent glossectomy and neck dissection and had histologically confirmed and clinically local or locoregional oral cavity cancer at the University of Texas MD Anderson Cancer Center were used.

For Human DRG experiments written informed consent for participation, including use of tissue samples, was obtained from each patient prior to inclusion. The protocol was reviewed and approved by The University of Texas MD Anderson Cancer Center Institutional Review Board, and all experiments conformed to relevant guidelines and regulations. Briefly, each donor was undergoing surgical treatment that necessitated ligation of spinal nerve roots to facilitate tumor resection or spinal reconstruction. Spinal roots were ligated proximal to the DRG, spinal roots were sharply cut both proximal and distal to the DRG, and excised DRG were transferred immediately into cold (~4°C) and sterile balanced salt solution containing nutrients. DRG were transported to the laboratory on ice in a sterile, sealed 50-mL centrifuge tube. Upon arrival to the laboratory, each ganglion was carefully dissected from the surrounding connective tissues and sectioned into several ~1- to 2-mm pieces. DRG were digested in 2 mL of a mixed enzyme solution: 0.1% trypsin (Sigma-Aldrich, T9201), 0.1% collagenase Sigma-Aldrich, C1764; w/v, final concentration), and 0.01% DNase (Sigma-Aldrich, D5025) diluted in DMEM/F-12. The pieces of tissue were transferred to a 37°C rotator to shake at a speed 124-128 revolutions/min. Every 20 minutes, tissue fragments were allowed to settle, and the supernatant/dissociated cells were collected and transferred to DMEM/F-12 with enzyme inhibitor. Supernatant was replaced with 2 mL of fresh digestion solution. The tissue was returned to the 37°C rotator, and this process was repeated until tissue fragments were well digested. Dissociated cells were centrifuged at 180 rpm for 5 minutes, supernatant was removed, and the cells were gently resuspended in culture medium with DMEM/F-12 supplemented with EV free 10% serum and 2 mM glutamine. Cells were plated onto laminin-coated μ -Slide 8 Well (ibidi) and cultured at 37°C with 5% CO₂ for 24-72 hours prior to undergoing transfection or labeling.

Recruitment

Human samples were obtained from surgically resected OCSG patients treated at the department of head and neck surgery at the University Texas MD Anderson Cancer Center, Houston, Tx USA. All patients gave their informed consent for the use of their resected specimen.

Ethics oversight

Study protocols were approved by the University of Texas MD Anderson Cancer Center Institutional Review Board.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

- ☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☐ All plots are contour plots with outliers or pseudocolor plots.
- ☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Primary sensory neurons were isolated from dissected TG as previously described (31). After ganglia dissection, tissue was enzymatically digested with papain (40 U/mL, EMD Millipore) for 20 minutes in 37°C followed by 20 minutes of digestion with collagenase II (4 mg/mL)/dispase II (4.6 mg/mL) solution. Using Percoll gradient (12.5% and 28% Percoll in complete L-15 medium [L-15 with 5% fetal calf serum, penicillin/streptomycin, HEPES]), we separated the myelin and nerve debris from trigeminal neurons. All cell sorting experiments (See Neuron Sequencing) were carried out using a FACSria Cell Sorter (BD Biosciences), and all flow cytometric analyses were carried out using an LSR II flow cytometer running FACSDiva 8.0 software (all BD Biosciences).

Samples were run in F-12 serum-free media, and cell doublets, debris, and dead cells were excluded by fetal calf serum, SSC, and NeuO (NeuroFluor NeuO, Stemcell Technologies, Ex/Em: 468/557 nm) profiles. Data were analyzed by Kaluza (Beckman Coulter) software. To assess adrenergic differentiation and transdifferentiation by flow cytometry, cell suspensions were stained on ice with cell surface markers and transcription factors (see Antibodies and staining reagents). Briefly, following cell fixation and permeabilization, cells were stained with the primary antibodies according to the manufacturer's instructions (BD Biosciences). After exclusion of dead cells and non-neuronal cells (by NTIII labeling), TH, OCT3/4, SOX2, KLF4, MASH1, doublecortin, TBR1 and neurogenin2 fluorescence was measured by flow cytometry.

Instrument

BD FACSria Cell Sorter and BD LSR II (BD Biosciences) were used in this study.

Software

BD FACS Diva Software Version 8.0.1 was used to collect the data.
FlowJo Version 1.0.1 was used to analyze the data.

Cell population abundance

Cell number subjected to flow cytometry was limited due to availability of the cells from trigeminal ganglia.

Gating strategy

FSC/SSC were used to discern single cells from doublets/multiple cells. Samples without fluorescent staining were used to establish boundaries between negative and positive cells.

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

SPEN integrates transcriptional and epigenetic control of X-inactivation

<https://doi.org/10.1038/s41586-020-1974-9>

Received: 4 June 2019

Accepted: 10 January 2020

Published online: 5 February 2020

François Dossin¹, Inês Pinheiro^{2,8}, Jan J. Żylicz^{2,3,8}, Julia Roensch², Samuel Collombet¹, Agnès Le Saux², Tomasz Chelmicki², Mikael Attia², Varun Kapoor², Ye Zhan⁴, Florent Dingli⁵, Damarys Loew⁶, Thomas Mercher⁶, Job Dekker^{4,7} & Edith Heard^{1,2*}

Xist represents a paradigm for the function of long non-coding RNA in epigenetic regulation, although how it mediates X-chromosome inactivation (XCI) remains largely unexplained. Several proteins that bind to *Xist* RNA have recently been identified, including the transcriptional repressor SPEN^{1–3}, the loss of which has been associated with deficient XCI at multiple loci^{2–6}. Here we show in mice that SPEN is a key orchestrator of XCI in vivo and we elucidate its mechanism of action. We show that SPEN is essential for initiating gene silencing on the X chromosome in preimplantation mouse embryos and in embryonic stem cells. SPEN is dispensable for maintenance of XCI in neural progenitors, although it significantly decreases the expression of genes that escape XCI. We show that SPEN is immediately recruited to the X chromosome upon the upregulation of *Xist*, and is targeted to enhancers and promoters of active genes. SPEN rapidly disengages from chromatin upon gene silencing, suggesting that active transcription is required to tether SPEN to chromatin. We define the SPOC domain as a major effector of the gene-silencing function of SPEN, and show that tethering SPOC to *Xist* RNA is sufficient to mediate gene silencing. We identify the protein partners of SPOC, including NCoR/SMRT, the m⁶A RNA methylation machinery, the NuRD complex, RNA polymerase II and factors involved in the regulation of transcription initiation and elongation. We propose that SPEN acts as a molecular integrator for the initiation of XCI, bridging *Xist* RNA with the transcription machinery—as well as with nucleosome remodellers and histone deacetylases—at active enhancers and promoters.

To assess the importance of SPEN during the initiation of XCI, we used an auxin-inducible degron (AID)⁷ that enables controlled and acute depletion of the endogenous SPEN protein. We used our previously described female hybrid (*Mus musculus castaneus* × C57BL/6) TX1072⁸ mouse embryonic stem cells (ES cells), in which a doxycycline (DOX)-inducible promoter upstream of the endogenous *Xist* locus enables conditional *Xist* RNA expression and XCI (Fig. 1a). In ES cells expressing the *Oryza sativa* TIR1 (OsTIR1) E3 ligase, we generated a homozygous knock-in that expressed the AID fused to a HaloTag at the C terminus of endogenous SPEN, in order to ensure auxin-dependent SPEN depletion (Extended Data Fig. 1a). Efficient degradation of SPEN occurred within 1 h of auxin treatment (Fig. 1b, Extended Data Fig. 1b, Supplementary Fig. 1) whereas the removal of auxin led to rapid recovery of SPEN (Fig. 1b), demonstrating potent AID-dependent modulation of SPEN levels.

To evaluate the immediate consequences of the loss of SPEN on the initiation of XCI, we acutely depleted SPEN for 4 h before inducing *Xist* expression for 24 h and performing RNA sequencing. Loss of SPEN had

no effect on the formation of *Xist* RNA clouds (Extended Data Fig. 1c, e), confirming that SPEN is dispensable for *Xist* localization^{2–5}. However, gene silencing was almost completely abolished along the entire X chromosome in the absence of SPEN (Fig. 1c, d, Supplementary Table 1), whereas auxin had no effect on XCI in wild-type cells (Extended Data Fig. 1d). Clustering analysis highlighted three groups of genes that differed in their silencing defects upon the loss of SPEN (Fig. 1e). Most X-linked genes (80% of 382) were found to be entirely dependent on SPEN for silencing, whereas only a small subset (6%) showed unaltered silencing in the absence of SPEN. This notable defect in XCI was confirmed by pyrosequencing (Fig. 1f) and nascent RNA fluorescence in situ hybridization (FISH) (Extended Data Fig. 1e).

We next assessed the requirement for SPEN in XCI in vivo during mouse early embryogenesis, using allele-specific RNA sequencing in embryonic day (E)3.5 *Spem*-knockout female embryos⁹ harbouring hybrid X chromosomes (Fig. 1g, Extended Data Fig. 1f, g). At this stage in wild-type embryos, imprinted XCI has taken place¹⁰ and only the paternal X chromosome is inactivated (Fig. 1h, Extended Data Fig. 1h).

¹European Molecular Biology Laboratory, Director's Unit, Heidelberg, Germany. ²Institut Curie, PSL Research University, CNRS UMR3215, INSERM U934, UPMC Paris-Sorbonne, Paris, France.

³Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK. ⁴Program in Systems Biology, Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA, USA. ⁵Institut Curie, PSL Research University, Centre de Recherche, Laboratoire de Spectrométrie de Masse Protéomique, Paris, France. ⁶INSERM U1170, Gustave Roussy Institute, Université Paris-Sud, Villejuif, France. ⁷Howard Hughes Medical Institute, Chevy Chase, MD, USA. ⁸These authors contributed equally: Inês Pinheiro, Jan J. Żylicz. *e-mail: edith.heard@embl.org

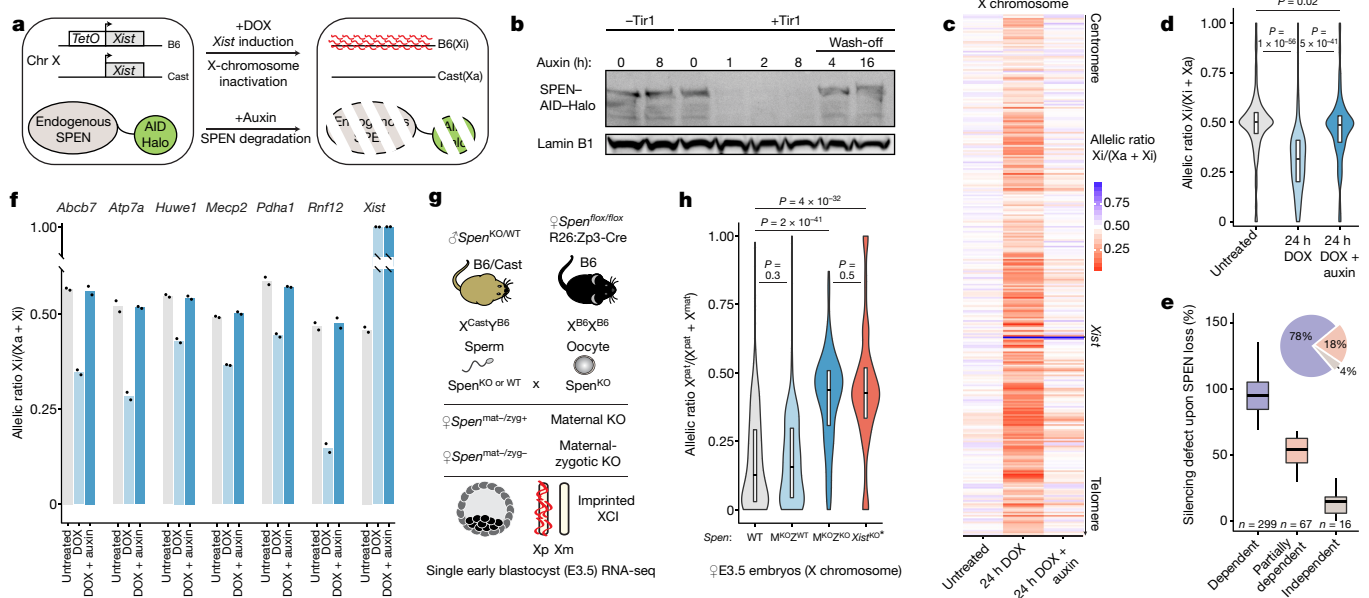


Fig. 1 | SPEN mediates gene silencing across the entire X chromosome in vitro and in vivo. **a**, Schematic of SPEN-degron *Xist*-inducible mouse ES cells. Xa, active X chromosome; Xi, inactive X chromosome. **b**, Western blot showing auxin-induced degradation of endogenous HaloTagged SPEN. This experiment was repeated at least twice with similar results. **c**, **d**, Heat map (**c**) and violin plots (**d**) showing X-chromosomal transcript allelic ratios after 0 h, 24 h DOX or 24 h DOX + auxin treatment in SPEN-degron mouse ES cells ($n = 434$ genes, two-sided Student's *t*-test). **e**, Box plot representation of gene-silencing defect upon SPEN loss in three groups of genes differing by their level of dependence on SPEN for *Xist*-mediated silencing. The pie chart shows the relative number of genes in each group. **f**, Pyrosequencing assay of seven X-linked transcripts in

mouse ES cells after 0 h, 24 h DOX or 24 h DOX + auxin treatment. Data in **c–f** are averages of two independent clones; in **f**, individual data points are shown. **g**, The mouse crossbreeding scheme for the *Spn*-knockout experiment. KO, knockout; WT, wild type; mat., maternal; zyg., zygotic; Xm, maternal X chromosome; Xp, paternal X chromosome. **h**, X-chromosomal transcript allelic ratio distribution ($n = 256$ genes) in wild-type ($n = 2$), maternal-only (*M*) *Spn*-knockout ($n = 3$), maternal-zygotic (*Z*) *Spn*-knockout ($n = 5$), and *Xist*-knockout E3.5 embryos ($n = 30$ single cells, two-sided Wilcoxon rank-sum test. For * refer to ref.¹⁰). In **d**, **e**, **h**, horizontal lines denote the median, box limits correspond to the upper and lower quartiles.

In maternal-zygotic *Spn* knockouts, imprinted XCI is severely hindered although paternal *Xist* is expressed. Both maternal and paternal X chromosomes are expressed equally, phenocopying *Xist*-knockout E3.5 embryos¹⁰ (Fig. 1h, Extended Data Fig. 1g, h, Supplementary Table 2). A maternal-only *Spn* knockout has no effect on imprinted XCI (Fig. 1h), suggesting that the zygotic pool of SPEN is necessary and sufficient for this process. Therefore, the early gene-silencing mechanism(s) involved in imprinted and random XCI are dependent on SPEN.

We next assessed precisely when SPEN is recruited during XCI. HaloTag labelling¹¹ of SPEN combined with *Xist* RNA FISH revealed that SPEN associates with *Xist* RNA rapidly upon *Xist* coating and throughout XCI (Fig. 2a). To capture early *Xist*–SPEN dynamics during the short time window in which *Xist* becomes upregulated, we followed both *Xist* and SPEN in living cells. We tagged endogenous SPEN with GFP in a background in which *Xist* RNA is visualized via a BglG–mCherry fusion protein binding to BglG stem-loops inserted within *Xist*¹² (Extended Data Fig. 2a, b). Live-cell imaging revealed that SPEN colocalizes with *Xist* from the very onset of *Xist* upregulation (Extended Data Fig. 2c, d, Supplementary Video 1). Therefore, SPEN can initiate gene silencing immediately upon *Xist* coating.

We also found that SPEN robustly accumulated on the inactive X chromosome after differentiation into neural progenitor cells (NPCs, Fig. 2b), in which XCI is epigenetically maintained. The depletion of SPEN for up to two days in independent NPC clones (Fig. 2c) did not lead to reactivation of fully silenced genes (Fig. 2d, Supplementary Table 3); however, we observed moderate but significant upregulation of genes escaping XCI (Fig. 2e, f), which suggests that SPEN buffers the overexpression of X-linked escapee genes in female cells.

Chromosome conformation capture has revealed that, in differentiated cells, the inactive X chromosome is folded into megadomains^{13–15}

and is globally depleted of topologically associating domains except in regions that contain clusters of escapee genes¹³. *Xist* RNA has been found to have a role in the conformation of the inactive X chromosome^{1,14}. To assess whether SPEN is involved, we performed allele-specific Hi-C in NPCs after 48 h of SPEN depletion. No notable conformational changes were observed on the inactive X chromosome (Extended Data Fig. 2e–g); we therefore conclude that the structural effects mediated by *Xist* RNA in differentiated cells occur independently of SPEN.

In summary, our data suggest that SPEN exerts its role by actively promoting gene silencing during the earliest stages of XCI. However, it has no major role in stabilizing the transcriptionally inactive state of the inactive X chromosome, or in ensuring the maintenance of its conformation.

We next sought to identify which parts of SPEN ensure its function during XCI. SPEN is a very large protein (around 400 kDa) that contains four RNA recognition motifs (RRMs), a nuclear receptor interaction domain (RID) and a SPEN paralogue/orthologue C-terminal (SPOC) domain (Fig. 3a). We overexpressed a series of SPEN complementary DNA truncations, stably targeted into the *Rosa26* locus in the SPEN-degron mouse ES cell line (Extended Data Fig. 3a, b, Fig. 3a). We then induced *Xist* expression for 24 h and assessed which SPEN fragments could rescue XCI-initiation function in the context of auxin-mediated depletion of endogenous SPEN. We found that the RRM1 domain and the RID are dispensable for SPEN accumulation on the inactive X chromosome, as well as for X-linked gene silencing (Fig. 3b, c). By contrast, a SPEN truncation lacking the RRM2–4 domains failed to accumulate on the inactive X chromosome and failed to rescue XCI (Fig. 3b, c). SPEN recruitment to the inactive X chromosome is therefore mediated by the RRM2–4 domains and is necessary for gene silencing. This is consistent with studies showing that RRM2–4 directly bind the A-repeat of *Xist*

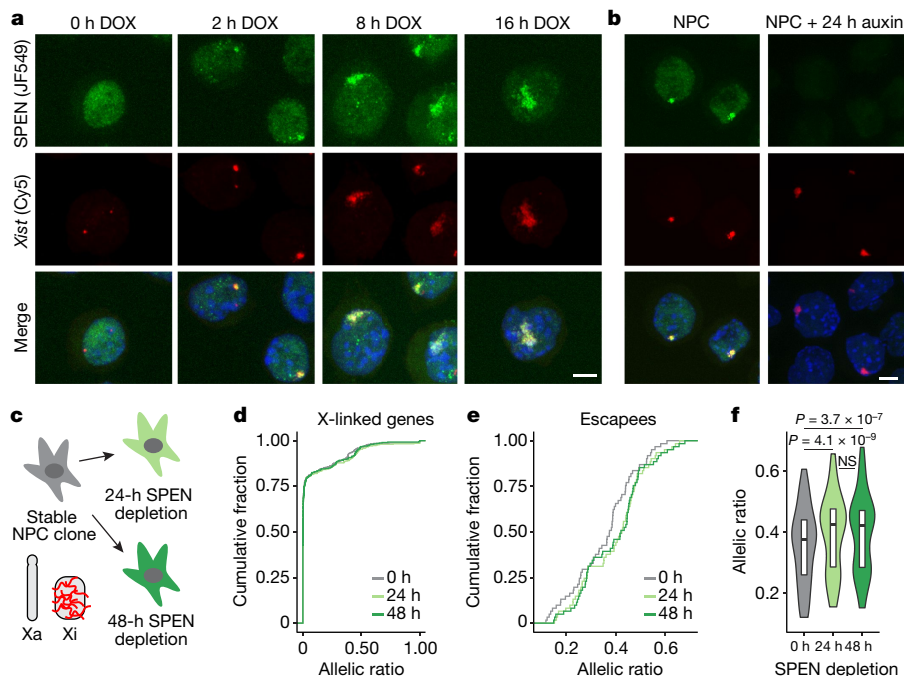


Fig. 2 | SPEN localizes to the X chromosome immediately upon *Xist* upregulation and throughout the stages of XCI, but is dispensable for the maintenance of X-linked gene silencing. **a, b,** Images from combined HaloTag labelling of SPEN (green) and FISH for *Xist* RNA (red) in mouse ES cells during a time course of *Xist* induction (**a**) and in NPCs (**b**). Scale bars, 5 μ m. **c,** Schematic of the SPEN-degron experiment in NPCs. **d,** Cumulative distribution of transcript allelic ratios across the X chromosome ($n = 387$ genes) after SPEN

depletion in NPCs. **e, f,** Cumulative distribution (**e**) and violin plot representation (**f**) of the transcript allelic ratio of escapees after SPEN depletion in NPCs ($n = 65$, two-sided Wilcoxon signed-rank test. NS, not significant. Horizontal lines denote the median, box limits correspond to upper and lower quartiles). Data in **d–f** are the average of two independent NPC clones. The experiments in **a, b** were repeated at least twice with similar results.

RNA in vitro^{4,16}—a region of *Xist* that is necessary for gene silencing¹⁷. Conversely, a truncation of the SPOC domain enabled efficient SPEN accumulation on the inactive X chromosome, but failed to rescue XCI (Fig. 3b, c). To validate this observation, we performed homozygous deletion of the SPOC domain at the endogenous *Spn* locus in mouse ES cells (Extended Data Fig. 3c). Deletion of the SPOC domain had no effect either on SPEN recruitment to the inactive X chromosome or on *Xist* RNA clouds (Extended Data Fig. 3d–f), but resulted in strongly deficient XCI, albeit milder than that in SPEN-depleted cells (Extended Data Fig. 3g–j). Collectively, these results demonstrate that the SPOC domain is essential for XCI. However, other uncharacterized regions of SPEN contribute—albeit to a lesser extent—to ensure its full silencing potential.

To test whether the SPOC domain alone could mediate X-linked gene silencing, we used SPEN-degron ES cells to introduce an array of Bgl stem-loops at the *Xist* locus (identical to the live-imaging strategy). In this background, we generated several independent ES cell lines expressing a BglG–GFP–SPOC protein fusion (or BglG–GFP as a control) targeted into *Rosa26*. These proteins would become tethered to *Xist*–Bgl stem-loop RNA via BglG (Fig. 3d). Notably, upon induction of *Xist* RNA in the absence of endogenous SPEN, tethering of BglG–GFP–SPOC (but not of BglG–GFP alone) resulted in substantial gene silencing across the X chromosome, with over half of the genes being silenced by more than 50% (Fig. 3e, f). SPOC-specific rescue was confirmed using pyrosequencing (Fig. 3g). Consistent with previous studies^{18–20}, our results reveal SPOC as a key domain of SPEN that enables gene silencing once recruited to the X chromosome by *Xist* RNA.

The SPOC domain of SPEN was originally identified as an interactor of the NCoR and SMRT corepressors in human cells^{18,21,22}. Given that NCoR and SMRT interact with and activate HDAC3²³, it was proposed that SPEN triggers XCI via HDAC3²³, the activity of which is important for *Xist*-mediated silencing^{2,24}. However, XCI is more markedly affected

upon the loss of SPEN and SPOC than upon the loss of HDAC3 (Extended Data Fig. 3j, k). These observations suggest that a model involving HDAC3 only partially explains the function of SPEN, and that SPOC must exert its key role in gene silencing also through other, HDAC3-independent pathways. To identify such pathways, we characterized the protein interactome of the SPOC domain by performing GFP pull-downs from mouse ES cells that stably expressed BglG–GFP–SPOC (or BglG–GFP as a control, Fig. 3h, Supplementary Table 4), followed by mass spectrometry analysis.

We identified NCoR and SMRT as expected, but we also found HDAC3 (Fig. 3h, Extended Data Fig. 3l), which further supports the proposed model for the function of SPEN in XCI². Notably, we identified the m⁶A methyltransferase complex and the m⁶A reader YTHDC1 (Fig. 3h, Extended Data Fig. 3l), which have been proposed to play a role in XCI^{5,20,25}. One of these factors, WTAP, co-purified with *Xist* RNA in an A-repeat-dependent manner³—although, contrary to the case of SPEN, a direct interaction between WTAP and *Xist* A-repeat has not been reported. Our results therefore suggest that SPOC may participate in the recruitment of m⁶A machinery to *Xist* RNA. We also identified the NuRD complex—a potent repressor that displaces RNA polymerase II (RNAPII) from transcription start sites through chromatin remodeling²⁶—and RNAPII, together with factors that are involved in the regulation of transcription initiation and elongation (Fig. 3h, Extended Data Fig. 3l). Together these findings show that, through its SPOC domain, SPEN bridges *Xist* to multiple factors that are involved in transcription and chromatin regulation, and together they mediate efficient gene silencing. Given that SPOC immunoprecipitation was performed in the absence of *Xist* induction, the identified interactions are not mediated by *Xist* RNA.

We also investigated where SPEN binds to the X chromosome during XCI, and whether it has distinct binding sites or whether it associates with chromatin diffusely across the entire chromosome, as anticipated

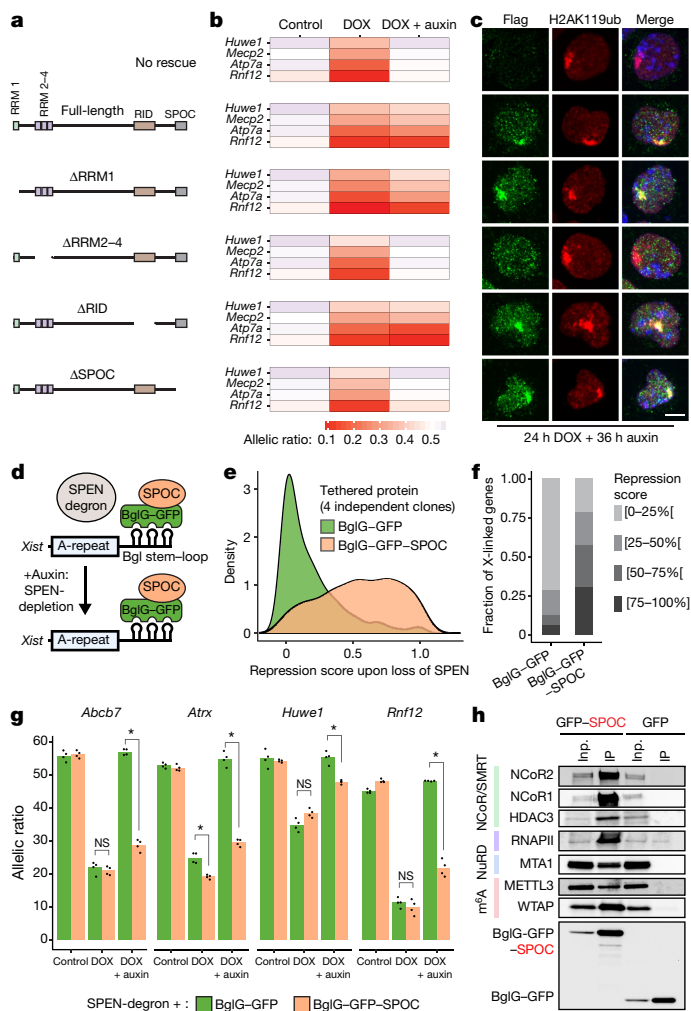


Fig. 3 | The SPOC domain of SPEN mediates gene silencing and interacts with multiple molecular pathways. **a**, *Spen* cDNA fragments used for the rescue experiment. **b**, Heat map representation of four X-linked transcript allelic ratios (obtained by pyrosequencing) in control, 24 h DOX- and 24 h DOX + 36 h auxin-treated SPEN-degron mouse ES cells overexpressing each cDNA construct. Data represent averages of two to three independent clones. **c**, Immunofluorescence detection of Flag-tagged SPEN truncations (green) and H2AK119ub1 (red), a marker of the inactive X chromosome, in SPEN-degron mouse ES cells treated with DOX and auxin. Scale bar, 5 μ m. **d**, Schematic showing the tethering of BglG-SPOC to *Xist*. **e**, Distribution of gene-repression scores observed across the X chromosome upon the depletion of endogenous SPEN and the tethering of BglG-GFP (green) or BglG-GFP-SPOC (orange) to *Xist*. **f**, Bar graphs showing the fraction of X-linked genes within four windows of repression score. **g**, Transcript allelic ratio (obtained by pyrosequencing) for four X-linked genes upon the depletion of endogenous SPEN and the tethering of BglG-GFP or BglG-GFP-SPOC to *Xist* (* $P < 0.01$, two-sided Student's *t*-test). **h**, Western blot showing co-immunoprecipitated proteins in BglG-GFP and BglG-GFP-SPOC immunoprecipitation experiments. One per cent of the input was loaded (0.1% for RNAPII), and 10% of the pull-down. The experiments in **c**, **h** were repeated at least twice with similar results. The data in **e–g** are the average of four independent clones.

from our imaging results. We performed allele-specific, cross-linked CUT&RUN²⁷ experiments on SPEN during a time course of *Xist* induction (0 h, 4 h, 8 h, 24 h DOX, or 8 h DOX + auxin as a negative control).

We found that there are few binding sites for SPEN across the genome of uninduced ES cells (Extended Data Fig. 4a). Conversely, hundreds of SPEN-binding sites appeared specifically on the X chromosome as early as 4 h after *Xist* induction (Fig. 4a, Extended Data Fig. 4a). This is consistent with imaging data (Extended Data Fig. 2). We note that SPEN

accumulation is seen across the gene body of *Xist* (Fig. 4b), suggesting that SPEN binds *Xist* RNA while it is transcribed. In sharp contrast to the *Xist* locus, SPEN shows focal binding on the rest of the genome, with peaks falling almost exclusively on promoters and enhancers (Fig. 4c, d, Extended Data Figs. 4b, 5a–g).

After *Xist* induction, recruitment of SPEN to the inactive X chromosome reaches a maximum at 4 h (Fig. 4a, Extended Data Fig. 4c), showing the highest enrichment within regions that were coated earliest by *Xist*²⁸ (entry sites, Fig. 4e). SPEN accumulation thus follows the spatial dynamics of *Xist* spreading. Among promoter targets on the X chromosome, SPEN preferentially binds those of actively expressed genes (Fig. 4f, Extended Data Fig. 4d), which suggests that the ability of SPEN to target chromatin depends on transcriptional activity. Consistently, genes that are classified as fully dependent on SPEN for silencing (Fig. 1e)—which show a greater degree of SPEN binding at their promoters within 4 h of *Xist* coating than less-dependent genes (Extended Data Fig. 4e)—also show initially higher transcription levels (Fig. 4g).

Furthermore, within 4 h of *Xist* induction, SPEN binding is greater at the promoters of efficiently silenced genes than at the promoters of less-efficiently silenced genes (Fig. 4h). Similarly, upon *Xist* coating, efficiently deacetylated enhancers²⁴ show a higher enrichment of SPEN than less-efficiently deacetylated enhancers (Fig. 4i). Finally, genes that are subject to very little silencing—or those that completely escape XCI in our *Xist*-inducible system—show a significantly lower SPEN signal at their promoters (Extended Data Figs. 4f, g, 5h–n). This pattern of SPEN recruitment at discrete sites to the X chromosome that is undergoing XCI indicates that transcriptional silencing is caused by the binding of SPEN to active promoters and enhancers.

To understand how SPEN might function at enhancers and promoters, we integrated CUT&RUN profiles with publicly available data from chromatin immunoprecipitation followed by sequencing (ChIP-seq) experiments for transcription and chromatin-associated factors identified in our mass spectrometry analysis. We included HDAC3²⁴, RNAPII²⁶ and two members of the NuRD complex (MBD3 and CHD4)²⁶. SPEN binding strongly overlaps with HDAC3 at enhancers but not at promoters (Extended Data Fig. 4h). Our recent findings revealed that HDAC3 is pre-bound predominantly at enhancers on the X chromosome²⁴. Therefore, *Xist*-mediated recruitment of SPEN to enhancers may activate HDAC3. Conversely, a strong overlap with SPEN binding is observed for the NuRD complex specifically at promoters but not at enhancers (Extended Data Fig. 4h). Furthermore, SPEN peaks extensively overlap with RNAPII phosphorylated on serine 5, which is associated with transcription initiation (Extended Data Fig. 4h). This analysis suggests that SPEN may operate at enhancers and promoters through distinct pathways to promote gene silencing.

Notably, the binding of SPEN to chromatin decreases across the whole X chromosome after 24 h of *Xist* induction (Fig. 4a, Extended Data Fig. 4c). Clustering of CUT&RUN profiles at SPEN-bound promoters (Extended Data Fig. 4i, Supplementary Table 5) revealed distinct groups of promoters, grouped on the basis of how efficiently SPEN was lost within 24 h of XCI (Fig. 4j). In the ‘strong SPEN loss’ group, binding was maximal by 4 h but decreased after 8 h, and even more markedly after 24 h (Fig. 4j). Conversely, the ‘mild SPEN loss’ group showed maximal and persistent SPEN binding at 4 h and 8 h of *Xist* induction, respectively, with only a mild reduction of SPEN binding by 24 h (Fig. 4j). Finally, a third group—comprising fewer promoters—showed both mild SPEN enrichment at 4 h and low loss at 24 h (Fig. 4j). The group that lost SPEN most efficiently also showed the most pronounced gene silencing by 24 h when compared with the groups that significantly retained SPEN (Fig. 4k, Extended Data Fig. 4j). Altogether this analysis suggests that, once recruited to the X chromosome by *Xist* RNA, SPEN associates with enhancers and promoters in a transcription-dependent manner. This recruitment leads to gene silencing, after which the favourable transcriptional context for SPEN binding is lost, and SPEN binding to chromatin decreases. Despite loss of the chromatin-bound SPEN

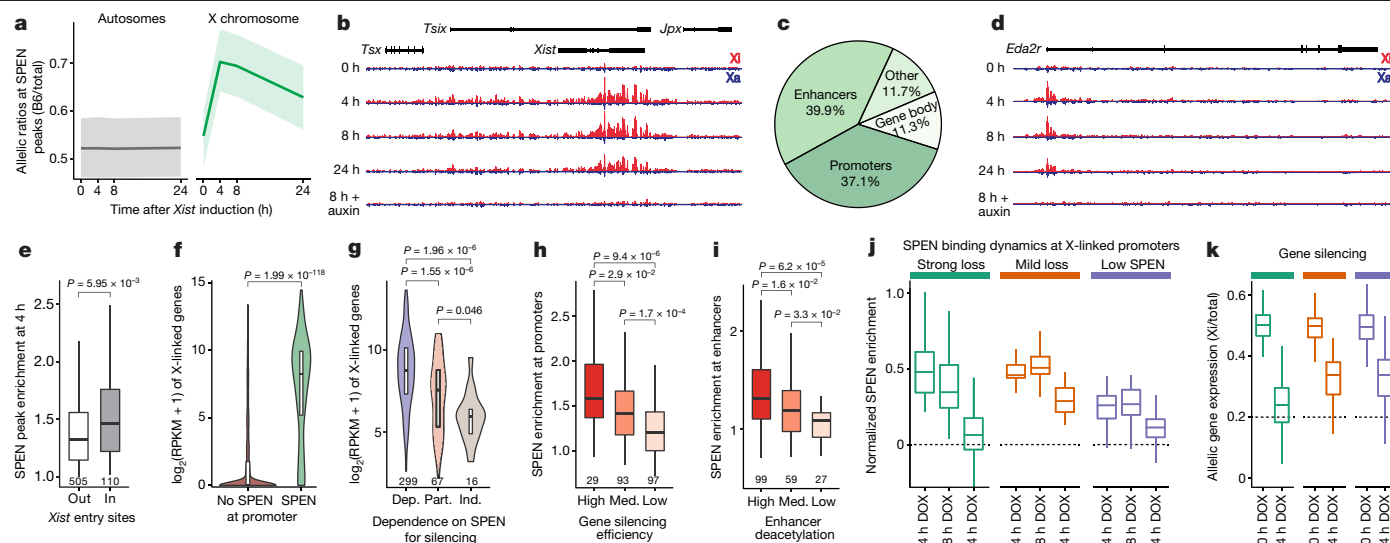


Fig. 4 | SPEN is recruited by *Xist* to active gene promoters and enhancers, where it silences transcription and subsequently disengages from chromatin. **a**, SPEN allele-specific accumulation (obtained from CUT&RUN experiments) on peaks at autosomes (grey, $n = 948$) and on the X chromosome (green, $n = 635$) after 0 h, 4 h, 8 h and 24 h of *Xist* induction in mouse ES cells. Shown are average allelic-ratios (shading is the interquartile range) of all peaks. **b**, UCSC Genome Browser allele-specific track showing SPEN binding around *Xist*. **c**, Annotation of SPEN peaks on the X chromosome. **d**, UCSC Genome Browser allele-specific track showing SPEN binding around *Eda2r*, an X-linked gene. In **b**, **d**, blue denotes Cast-Xa; red denotes B6-Xi; tracks are scaled identically. **e**, Box plot showing SPEN enrichment at 4 h in peaks outside or within *Xist* entry sites. **f**, Violin plot showing gene expression (reads per kilobase per million reads, RPKM) of genes accumulating SPEN ($n = 289$) or not

accumulating SPEN ($n = 2,325$) at their promoters. **g**, Violin plot showing gene expression levels (RPKM in control conditions) of genes grouped on the basis of their level of dependence on SPEN for gene silencing (see Fig. 1e). **h**, **i**, Box plots showing SPEN enrichment after 4 h of *Xist* induction within peaks at promoters grouped on the basis of how efficiently their respective genes are silenced (**h**) or at enhancers grouped on the basis of how efficiently they are deacetylated during XCI (**i**). **j**, **k**, Box plots showing normalized SPEN enrichment at promoters (**j**) and gene silencing (transcript allelic ratio) during XCI (**k**) within 3 groups of X-linked genes showing different dynamics of SPEN accumulation and loss ($n = 86$ strong loss, $n = 92$ mild loss, $n = 39$ low SPEN). In **e**–**i**, the two-sided Wilcoxon rank-sum test was used; in **e**–**k**, horizontal lines denote the median, box limits correspond to upper and lower quartiles.

fraction, persistent *Xist* RNA expression and coating ensure that SPEN remains strongly accumulated around the inactive X chromosome (Fig. 2a, b).

Our study demonstrates that SPEN is a crucial factor that collaborates with *Xist* RNA to initiate gene silencing across the X chromosome, both during XCI in vitro and imprinted XCI in vivo. SPEN becomes dispensable for maintaining gene silencing after XCI has been established, but partially represses escapees, which suggests that *Xist* may have a silencing role even in somatic cells. Although SPEN coats the X chromosome immediately upon *Xist* induction, it contacts chromatin only at active promoters and enhancers, which serve as substrates for SPEN-mediated gene silencing. SPEN association with chromatin is favoured by active transcription, as SPEN disengages from chromatin when X-linked genes become silenced. We identify the SPOC domain of SPEN as a potent transcriptional repressor, which is crucial for SPEN-dependent XCI. On the basis of our mass spectrometry analysis, we propose that the SPOC domain is key for bridging *Xist* with other factors implicated in XCI—such as HDAC3—which we find to be present at most X-linked enhancers to which SPEN is recruited. In particular, the interaction of the SPOC domain with the NuRD complex and the transcription machinery points to a role for SPEN in direct transcriptional repression. We also identify SPOC as an interactor of the m⁶A methyltransferase complex, which has a role in *Xist* RNA methylation, a modification that is important for *Xist*-dependent silencing²⁵. Methylation of *Xist* is mediated by RBM15²⁵, which interacts with the m⁶A machinery directly through ZC3H13²⁹—the most highly enriched m⁶A machinery factor identified in our mass spectrometry experiments. Because RBM15 also carries a SPOC domain, our study raises the possibility that the interaction with the RNA methylation machinery is not restricted solely to the SPOC domain of SPEN, but may instead be a feature that is shared across SPOC-containing proteins.

SPEN binds other non-coding RNAs, including *SRA*¹⁸, which is involved in steroid-receptor regulation. Furthermore, another *SRA*-binding protein—SLIRP—has been shown to bind promoters in an *SRA*-dependent manner³⁰; this raises the possibility that, similarly to *Xist*, *SRA* could guide SPEN to target gene regulatory elements.

In conclusion, our study suggests that RNA-mediated recruitment of SPEN and other SPOC-containing proteins—which are found across fungi, plants and animals—may be a widespread means by which to acutely repress transcription by co-ordinately engaging several layers of epigenetic and transcriptional control. We propose that SPEN bridges *Xist* to the transcription machinery, histone deacetylases and chromosome remodelling factors to ensure robust and efficient XCI (Extended Data Fig. 4k).

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-1974-9>.

- Minajigi, A. et al. A comprehensive *Xist* interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* **349**, aab2276 (2015).
- McHugh, C. A. et al. The *Xist* lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **521**, 232–236 (2015).
- Chu, C. et al. Systematic discovery of *Xist* RNA binding proteins. *Cell* **161**, 404–416 (2015).
- Monfort, A. et al. Identification of *Spn* as a crucial factor for *Xist* function through forward genetic screening in haploid embryonic stem cells. *Cell Rep.* **12**, 554–561 (2015).
- Moindrot, B. et al. A pooled shRNA screen identifies Rbm15, Spn, and Wtap as factors required for *Xist* RNA-mediated silencing. *Cell Rep.* **12**, 562–572 (2015).
- Nesterova, T. B. et al. Systematic allelic analysis defines the interplay of key pathways in X chromosome inactivation. *Nat. Commun.* **10**, 3129 (2019).

7. Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T. & Kanemaki, M. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat. Methods* **6**, 917–922 (2009).
8. Schulz, E. G. et al. The two active X chromosomes in female ESCs block exit from the pluripotent state by modulating the ESC signaling network. *Cell Stem Cell* **14**, 203–216 (2014).
9. Yabe, D. et al. Generation of a conditional knockout allele for mammalian Spen protein Mint/SHARP. *Genesis* **45**, 300–306 (2007).
10. Borensztein, M. et al. *Xist*-dependent imprinted X inactivation and the early developmental consequences of its failure. *Nat. Struct. Mol. Biol.* **24**, 226–233 (2017).
11. Grimm, J. B. et al. A general method to improve fluorophores for live-cell and single-molecule microscopy. *Nat. Methods* **12**, 244–250 (2015).
12. Masui, O., Heard, E. & Koseki, H. in *X-Chromosome Inactivation* (ed. Sado, T.) *Methods Mol. Biol.* Vol. 1861, 67–72 (Humana, 2018).
13. Giorgetti, L. et al. Structural organization of the inactive X chromosome in the mouse. *Nature* **535**, 575–579 (2016).
14. Deng, X. et al. Bipartite structure of the inactive mouse X chromosome. *Genome Biol.* **16**, 152 (2015).
15. Rao, S. S. P. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).
16. Lu, Z. et al. RNA duplex map in living cells reveals higher-order transcriptome structure. *Cell* **165**, 1267–1279 (2016).
17. Wutz, A., Rasmussen, T. P. & Jaenisch, R. Chromosomal silencing and localization are mediated by different domains of *Xist* RNA. *Nat. Genet.* **30**, 167–174 (2002).
18. Shi, Y. et al. Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. *Genes Dev.* **15**, 1140–1151 (2001).
19. Oswald, F. et al. RBP-Jk/SHARP recruits CtIP/CtBP corepressors to silence Notch target genes. *Mol. Cell. Biol.* **25**, 10379–10390 (2005).
20. Ha, N. et al. Live-cell imaging and functional dissection of *Xist* RNA reveal mechanisms of X chromosome inactivation and reactivation. *iScience* **8**, 1–14 (2018).
21. Ariyoshi, M. & Schwabe, J. W. R. A conserved structural motif reveals the essential transcriptional repression function of Spen proteins and their role in developmental signaling. *Genes Dev.* **17**, 1909–1920 (2003).
22. Oswald, F. et al. A phospho-dependent mechanism involving NCoR and KMT2D controls a permissive chromatin state at Notch target genes. *Nucleic Acids Res.* **44**, 4703–4720 (2016).
23. Guenther, M. G., Barak, O. & Lazar, M. A. The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3. *Mol. Cell. Biol.* **21**, 6091–6101 (2001).
24. Żylicz, J. J. et al. The implication of early chromatin changes in X chromosome inactivation. *Cell* **176**, 182–197 (2019).
25. Patil, D. P. et al. m⁶A RNA methylation promotes *XIST*-mediated transcriptional repression. *Nature* **537**, 369–373 (2016).
26. Bornelöv, S. et al. The nucleosome remodeling and deacetylation complex modulates chromatin structure at sites of active transcription to fine-tune gene expression. *Mol. Cell* **71**, 56–72 (2018).
27. Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife* **6**, e21856 (2017).
28. Engreitz, J. M. et al. The *Xist* lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* **341**, 1237973 (2013).
29. Knuckles, P. et al. Zc3h13/Flacc is required for adenosine methylation by bridging the mRNA-binding factor Rbm15/Spenito to the m⁶A machinery component Wtap/Fl(2)d. *Genes Dev.* **32**, 415–429 (2018).
30. Hatchell, E. C. et al. SLIRP, a small SRA binding protein, is a nuclear receptor corepressor. *Mol. Cell* **22**, 657–668 (2006).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Methods

Data reporting and statistical analysis

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment. All statistical tests, resulting *P* values and observation numbers are indicated in the figure panels or in the figure legends.

Data visualization

All heat maps, violin plots, box plots, density plots, bar graphs and pie charts were generated using ggplot2. Unless stated otherwise, box plots always show the median as the centre line, box limits correspond to upper and lower quartiles, and whiskers cover 1.5× the interquartile range.

Plasmid construction

The plasmids to target *OsTIR1* at the *TIGRE* locus (Addgene plasmid 92141) and the *TIGRE*-specific guide-RNA-encoding plasmid (Addgene plasmid 92144) were provided by E. Nora. The additional *TIGRE*-targeting plasmids BglG-mCherry-T2A-*OsTir1* (pFD51) and rtTa-VP16-T2A-*OsTir1* (pFD68) were cloned using PCR amplification of corresponding gene cassettes followed by traditional cloning into the 92141 backbone.

Targeting constructs (pFD19 and pFD49) to tag endogenous SPEN at its C terminus with AID-HaloTag and AID-GFP, respectively, were generated as follows: 500-bp homology arms (flanking both sides of, but excluding the stop codon of *Spn*) were amplified from mouse genomic DNA by PCR. One-step Gibson cloning (New England Biolabs) was subsequently used to simultaneously surround the digested AID insert (carrying a puromycin-resistance gene under the control of the PGK promoter) in frame with the homology arms and clone the insert into the pBR322 vector. Synonymous mutations in the PAM/SEED target sequence (located on the 5' homology arm) were then introduced using the QuickChange II XL site-directed mutagenesis kit (Agilent) to prevent Cas9-mediated cutting of the targeting vector upon transfection and of the AID-tagged allele(s) upon integration. The targeting construct (pFD90) to replace the endogenous SPOC domain of SPEN by GFP was generated using the same strategy. For guide RNA cloning, the pX459 plasmid (a gift from F. Zhang, Addgene 62988) encoding *Streptococcus pyogenes* Cas9 was digested with BbsI immediately downstream of the U6 promoter, and annealed DNA duplexes corresponding to the target guide RNA sequences were ligated.

Cell culture

Mouse XX ES cells (TX1072) were grown on 0.1% gelatin-coated flasks in 8% CO₂ 37 °C incubators. For all experiments, cells were cultured in 2i+LIF, and batch-tested fetal calf serum ES cell medium – DMEM (Sigma), 15% FBS (Gibco), 0.1 mM β-mercaptoethanol, 1,000 U ml⁻¹ leukaemia inhibitory factor (LIF, Chemicon), CHIR99021 (3 μM), PD0325901 (1 μM).

NPC differentiations and subcloning were performed as previously described¹³. NPCs were grown in N2B27 medium supplemented with EGF and FGF (10 ng ml⁻¹ each), on 0.1% gelatin-coated flasks.

Cell transfection and clone isolation

All transgenic insertions were performed using the 4D nucleofector system from Lonza. For each nucleofection, five million cells were electroporated with 2.5 μg each of non-linearized targeting vectors and guide RNA-Cas9 encoding plasmids (MidiPreps). Nucleofected cells were then serially diluted and plated on 10-cm dishes. Forty-eight hours later, antibiotic selection was performed (puromycin, 0.4 μg ml⁻¹; hygromycin, 250 μg ml⁻¹; blasticidin, 5 μg ml⁻¹), except for transfection steps involving flippase-mediated removal of resistance cassettes, during which no selection was applied. One week after the initial plating, 80 to 96 single colonies were picked from dishes showing ideal

clonal density and seeded in 96-well plates. These cells were subsequently split into one high-confluency plate used for PCR genotyping, and one low-confluency plate from which desired clones were further expanded until T25 density was reached. At this stage, some cells were kept to reconfirm the correct genotype by PCR, while the remaining cells were frozen.

Cell treatments

Xist expression in TX1072 mouse ES cells was induced upon administration of doxycycline (1 μg ml⁻¹). Auxin-mediated depletion of target proteins was achieved by supplementing culture media with auxin (Sigma) at the recommended concentration of 500 μM. Auxin-containing medium was renewed every 24 h. For auxin wash-out, auxin-containing medium was removed, cells were rinsed once with PBS, and exposed to auxin-free medium.

Protein extraction and western blotting

Cells were trypsinized, washed once in medium and once in PBS and then pellets were immediately frozen at –80 °C. Pellets were then resuspended in RIPA buffer (50 mM Tris-HCl pH 8.0–8.5, 150 mM NaCl, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS) containing protease inhibitors (Roche), incubated for 30 min on ice and sonicated with a Bioruptor (three 10-s pulses). Lysates were then centrifuged for 20 min at 4 °C, and supernatants were kept. Protein concentration was determined using the Bradford (BioRad) assay. Samples were then boiled at 95 °C for 10 min in LDS buffer (Thermo) containing 200 mM DTT. For all western blots except those aimed at detecting SPEN, 4–12% Bis-Tris gels were used. For the detection of SPEN, a high-molecular-weight protein (>400 kDa), 3–8% tris-acetate polyacrylamide gels were used. Transfer was performed on a 0.45-μm nitrocellulose membrane using a wet-transfer system, at 350–400 mA for 2 h at 4 °C.

RNA extraction, reverse transcription, pyrosequencing and RNA sequencing

RNA extraction was performed using the RNeasy kit and on-column DNase digestion (Qiagen). Reverse transcription was performed on 1 μg total RNA using SuperScript III (Life Technologies). To quantify allelic skewing, cDNA was amplified using biotinylated primers and subsequently sequenced using Q24 Pyromark (Qiagen). Only samples showing a RNA integrity number greater than 9 were used to prepare RNA sequencing (RNA-seq) libraries (TruSeq). Paired-end 100-nt sequencing was performed on a HiSeq2500 or NovaSeq6000 (Illumina).

RNA FISH

Cells were dissociated using Trypsin (Invitrogen) for ES cells or Accutase (Invitrogen) for NPCs, washed twice in medium, and allowed to attach on poly-L-lysine (Sigma)-coated coverslips for 10 min. Cells were fixed with 3% paraformaldehyde in PBS for 10 min at room temperature, washed in PBS three times, and permeabilized with ice-cold permeabilization buffer (PBS, 0.5% Triton X-100, 2 mM vanadyl-ribonucleoside complex) for 5 min on ice. Coverslips were stored in 70% ethanol at –20 °C. Samples were dehydrated in 4 baths of increasing ethanol concentration (80%, 95%, 100% twice) and air-dried quickly. Probes were prepared from minipreps of intron-spanning bacteria artificial chromosomes (BACs) (clone RP24-157H12 for *Huwei1*, RP23-260I15 for *Atrx*) or plasmid (p510 for *Xist*). Probes were labelled by nick translation (Abbott) using dUTP labelled with spectrum green (Abbott) for *Huwei1*, spectrum red (Abbott) for *Atrx*, and Cy5 (Merck) for *Xist*. Labelled BAC probes were co-precipitated with Cot-1 DNA repeats in the presence of ethanol and salt, resuspended in formamide, denatured at 75 °C for 10 min, and competed at 37 °C for 1 h. Probes were then co-hybridized in FISH hybridization buffer (50% formamide, 20% dextran sulfate, 2X SSC, 1 μg μl⁻¹ BSA, 10 mM vanadyl-ribonucleoside) at 37 °C overnight. The next day, hybridized coverslips were washed three times for 5 min

Article

with 50% formamide in 2X SSC at 42 °C, and three times for 5 min with 2X SSC. DAPI (0.2 mg ml⁻¹) was added to the penultimate wash and coverslips were mounted with Vectashield (Vectorlabs).

HaloTag labelling

HaloTag labelling of the SPEN–Halo fusion protein was performed in live TX1072 ES cells and NPCs. Cells were labelled with HaloTag–ligand-conjugated Janelia Fluor¹¹ (JF646–HaloTag or JF549–HaloTag, a gift from L. Lavis) at a final concentration of 250 nM in culture medium. Labelling was performed for 1 h at 37 °C, cells were then washed 4 times with generous volumes of PBS, and incubated with unlabelled medium for 15 min before proceeding with downstream experiments. For NPC labelling, cells were washed with unlabelled medium and not PBS, because NPCs detach when exposed to PBS. Auxin and/or doxycycline were kept in the labelling medium when necessary.

HaloTag labelling followed by RNA FISH

For co-detection of SPEN–Halo and *Xist* RNA, cells were labelled with JF549 as indicated above, and directly processed for fixation and permeabilization as detailed in the section ‘RNA FISH’. Importantly, after permeabilization, coverslips were directly washed twice with PBS, twice with 2X SSC and immediately processed for FISH.

Mouse breeding, embryo collection and single-embryo RNA-seq

Timed natural matings were used for all experiments. Noon of the day when the vaginal plugs of mated females were identified was scored as E0.5. For *Spn* matings a conditional allele was used⁹. For oocyte deletions the published *Rosa26:Zp3-Cre* allele was used³¹. F₁ hybrid *Spn*^{+/-} males were obtained by crossing *Spn*^{+/-} CAST/Eij females with *Spn*^{+/-} C57BL/6J males. For *Spn* maternally deleted embryos, *Spn*^{flx/flx} *Zp3-Cre*^{+/ve} C57BL/6J females were crossed with *Spn*^{+/-} F₁ hybrid males. For *Spn* control embryos, *Spn*^{flx/flx} *Zp3-Cre*^{-ve} C57BL/6J females were crossed with *Spn*^{+/-} F₁ hybrid males. The care and use of animals in this study was performed in accordance with the recommendations of the European Community (2010/63/UE). All experimental protocols were approved by the ethics committee of Institut Curie CEEA-IC118 under the number APAFIS#8812-2017020611033784v2 given by national authority in compliance with the international guidelines. Single-embryo RNA-seq was performed as previously described³². In brief, E3.5 embryos were collected and morphologically assessed to ensure that only viable samples were collected. The zona pellucida was removed by treatment with acidified Tyrode’s solution. Single embryos were picked into individual tubes and cDNA was prepared and amplified as previously described³³. Illumina libraries were prepared as published in ref. ³⁴. Paired-end 100-nt sequencing was performed with HiSeq2500 (Illumina).

Live-cell imaging and analysis

Cells were seeded on fibronectin-coated 35-mm glass-bottom dishes (Ibidi) 24 h before imaging. Doxycycline was added 1 h before image acquisition. Cells were imaged on the DeltaVision OMX microscope in widefield mode (GE Healthcare) using a 1.4 numerical aperture 100× oil immersion objective. The temperature was controlled at 37 °C and CO₂ at 8% during acquisition.

Images were acquired as z-stacks of 40 slices with 400-nm steps every 10 min for at least 4 h. Movies were deconvolved using Huygens deconvolution with the following parameters: Iteration 4; S/N 5, 10; quality threshold 0.1; and widefield mode 0.7 was used for background estimation. Two channels were registered using TetraSpec microspheres 0.1 μm (Invitrogen) and unwarped (Fiji plug-in). For segmentation, z-projected deconvolved registered images were used and pixels were classified as cloud or nuclei using Ilastik. Touching nuclei were sometimes manually separated. Cut-offs on resulting probability maps were set to 0.7. We next performed connected component analysis to obtain integer-labelled images in which each integer label corresponds

to a unique nucleus. In the tailor-made Fiji plug-in the inputs are the raw max z-projected time-lapse images of the two channels and the integer labelled time-lapse image of the nuclei. The probability maps of the clouds give the region of interest in the time-lapse sequence in which total intensity is calculated. In the plug-in, clouds are associated with their corresponding nuclei; they are then linked via Kalman filter tracker over time. These unique links constitute track IDs and contain information about the intensity and area measurements for each cell. For each tracked cell, the first time point when a cloud is detected in one channel (*Xist* or SPEN) is labelled as reference time point 1.

Hi-C

Hi-C was performed as previously described³⁵, except that ligated DNA size selection was omitted, and dA-tailing was performed before biotin pull-down. In brief, each Hi-C experiment was performed on 10 million cells (NPCs) per sample. Cells were digested with DpnII at 37 °C overnight. DNA ends were filled with biotin-14-dATP at 23 °C for 4 h. DNA was then ligated with T4 DNA ligase at 16 °C overnight. Binding proteins were removed by treating ligated DNA with proteinase K at 65 °C overnight. Purified proximally ligated molecules were fragmented to obtain an average fragment size of 200 bp. After DNA end repair, dA-tailing and biotin enrichment, DNA molecules were ligated to Illumina TruSeq sequencing adapters at room temperature for 2 h. Final library PCR productions were carried out following the Illumina TruSeq Nano DNA Sample Prep Kit manual. Paired-end 100-nt sequencing was performed on a HiSeq4000 (Illumina).

Genetic engineering strategy for *Xist*–Bgl stem-loop tagging and SPEN complementation analysis constructs

To tag *Xist* with Bgl stem-loops³⁶, we nucleofected cells with pBS-Ptight-*Xist*-BglSL¹² (plasmid harbouring 18 repeats of Bgl stem-loops inserted between homology arms to target *Xist* exon 7, carrying a G418 selection gene, a gift from O. Masui). After G418 selection and FLP-FRT mediated removal of the selection cassette, clones were picked and genotyped. Positive clones were further tested to ensure that the stem-loop-tagged *Xist* could properly be induced and trigger gene silencing upon addition of doxycycline (data not shown).

Spn cDNA truncations were generated by splicing out different regions of the *Spn* open reading frame (Genscript, ORF clone OMu11416C) using overlap extension PCR. Each *Spn* truncation was cloned downstream of a CAGGS promoter into a vector carrying homology arms for targeted insertion at the *Rosa26*³⁷ locus as well as a SV40-promoter driven hygromycin-resistance gene. The BglG–GFP–SPOC targeting plasmid was designed by inserting a translational fusion between a BglG–GFP cassette and SPEN amino acids 3244–3643 into the same *Rosa26* targeting vector. Each of these ‘complementation’ constructs were independently targeted at *Rosa26* in SPEN-degrogen mouse ES cells. Independent clones were picked and protein expression of each SPEN truncation was assessed by western blot. XCI complementation analysis was then performed in 2–3 independent clones for *Spn* cDNA truncations, and 4 independent clones for BglG–GFP–SPOC expressing clones. The ability of cells to accumulate BglG–mCherry, BglG–GFP and BglG–GFP–SPOC upon the addition of doxycycline was assessed using microscopy (data not shown).

Immunofluorescence

ES cells were dissociated using trypsin (Invitrogen), washed extensively in medium, and allowed to attach on poly-L-lysine (Sigma)-coated coverslips for 10 min. Cells were then fixed with 3% paraformaldehyde in PBS for 10 min at room temperature, washed in PBS three times, and permeabilized with 0.25% Triton X-100 in PBS for 5 min at room temperature. Coverslips were then washed three times in PBS and blocked for 1 h with blocking buffer (PBS containing 2.5% BSA, 0.1% Tween20 and 10% normal goat serum). Coverslips were then incubated with primary antibodies diluted in blocking buffer at 4 °C overnight, washed three

times for 5 min in PBST (0.1% Tween20) the next day, incubated with fluorescently labelled secondary antibodies (1/500 in blocking buffer) for 1 h at room temperature, and washed again three times for 5 min in PBST. DAPI (0.2 mg ml⁻¹) was added to the penultimate wash and coverslips were mounted with Vectashield (Vectorlabs).

Immunoprecipitation

Nuclear extracts were prepared by resuspending 50 million fresh cells in ice-cold 10 ml buffer A (10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM MgCl₂, 0.1% NP-40, cOmplete EDTA free, phosSTOP) and rotating for 10 min at 4 °C. Nuclei were centrifuged at 800g for 10 min at 4 °C and resuspended in 1 ml IP buffer C150 (20 mM HEPES pH 7.9, 150 mM NaCl, 1.5 mM MgCl₂, 0.2 mM EDTA, 0.25% NP-40, cOmplete EDTA free, phosSTOP). Lysates were briefly sonicated followed by Benzonase (Merck) digestion for 30 min at 4 °C. Finally, lysates were cleared through centrifugation at 13,000 rpm for 20 min before being incubated with 15 µl of GFP trap magnetic agarose bead slurry (ChromoTek) overnight at 4 °C. Beads were washed 5 times in IP buffer. For co-immunoprecipitation (Co-IP) western blot, washed beads were directly resuspended in LDS buffer (Thermo) containing 200 mM DTT, and boiled at 95 °C for 10 min.

Proteomics and mass spectrometry analysis

Proteins on magnetic beads were washed twice with 100 µl of 25 mM NH₄HCO₃ and we performed on-bead digestion with 0.2 µg of trypsin/LysC (Promega) for 1 h in 100 µl of 25 mM NH₄HCO₃. Samples were then loaded onto homemade C18 StageTips for desalting. Peptides were eluted using 40/60 MeCN/H₂O + 0.1% formic acid and concentrated to dryness under vacuum. Online chromatography was performed with an RSLCnano system (Ultimate 3000, Thermo Scientific) coupled online to a Q Exactive HF-X with a Nanospray Flex ion source (Thermo Scientific). Peptides were first trapped on a C18 column (75 µm inner diameter × 2 cm; nanoViper Acclaim PepMap 100, Thermo Scientific) with buffer A (2/98 MeCN/H₂O in 0.1% formic acid) at a flow rate of 2.5 µl min⁻¹ over 4 min. Separation was then performed on a 50 cm × 75 µm C18 column (nanoViper Acclaim PepMap RSLC, 2 µm, 100 Å, Thermo Scientific) regulated to a temperature of 50 °C with a linear gradient of 2% to 30% buffer B (100% MeCN in 0.1% formic acid) at a flow rate of 300 nl min⁻¹ over 91 min. Mass spectrometry full scans were performed in the ultra-high-field Orbitrap mass analyser over the range *m/z* 375–1,500 with a resolution of 120,000 at *m/z* 200. The top 20 most intense ions were subjected to Orbitrap for further fragmentation via high-energy collision dissociation activation and a resolution of 15,000 with the intensity threshold kept at 1.3×10^5 . We selected ions with charge state from 2⁺ to 6⁺ for screening. Normalized collision energy was set at 27 and the dynamic exclusion at 40 s. For identification, the data were searched against the *M. musculus* (UP000000589) Uniprot database using Sequest HF through Proteome Discoverer (v.2.2). Enzyme specificity was set to trypsin and a maximum of two missed-cleavage sites were allowed. Oxidized methionine and N-terminal acetylation were set as variable modifications. Maximum allowed mass deviation was set to 10 ppm for monoisotopic precursor ions and 0.02 Da for MS/MS peaks. The resulting files were further processed using myProMS³⁸ v3.6 (work in progress). Calculation of the false discovery rate used Percolator and was set to 1% at the peptide level for the whole study. The label-free quantification was performed by peptide extracted ion chromatograms (XICs) computed with MassChroQ version 2.2³⁹. For protein quantification, XICs from proteotypic peptides shared between compared conditions (TopN) with two missed cleavages were used. Median and scale normalization was applied on the total signal to correct the XICs for each biological replicate. To estimate the significance of the change in protein abundance, a linear model (adjusted on peptides and biological replicates) was performed and *P* values were adjusted with a Benjamini–Hochberg false-discovery rate procedure with a control threshold set to 0.05. The mass spectrometry proteomics

data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD015699.

Cross-linked CUT&RUN

CUT&RUN against SPEN was performed during a timecourse of *Xist* induction/SPEN degradation: 0 h DOX, 4 h DOX, 8 h DOX, 24 h DOX and 8 h DOX + auxin. Two biological replicates were performed. The original CUT&RUN protocol²⁷ was adapted for fixed cells: 10⁶ cells in suspension were fixed with 2% formaldehyde diluted in PBS for 10 min at room temperature (2 ml final volume). Fixation was quenched with 125 mM glycine for 5 min and cells were washed twice in 1 ml PBS. Fixed cells were then permeabilized with 1 ml permeabilization buffer (20 mM HEPES pH 7.9, 150 mM NaCl, 0.5 mM spermidine, 0.25% TritonX-100, cOmplete EDTA free) for 5 min and washed twice in 1 ml PBS. Cells were then resuspended in 1 ml washing buffer (20 mM HEPES pH 7.9, 150 mM NaCl, 0.5 mM spermidine, 0.1% BSA, cOmplete EDTA free), bound to activated concanavalin beads (50 µl bead slurry used per 10 million cells) for 10 min, and blocked in 1 ml blocking buffer (wash buffer + 2 mM EDTA) for 5 min. At this stage, cells were resuspended in 500 µl wash buffer containing target antibodies diluted 1/200, transferred to 0.5-ml tubes, and incubated overnight at 4 °C on an end-to-end rotator. Cells were washed three times in 500 µl washing buffer followed by 1-h incubation with pA-MNase (500 µl of washing buffer containing 700 ng ml⁻¹ pA-MNase, produced by the Protein Expression and Purification Core Facility of Institut Curie) and washed again three times in 500 µl washing buffer. After the last wash, cells were resuspended in 150 µl washing buffer, transferred to 1.5-ml tubes, and equilibrated to 0 °C in a metal block for 10 min. To start digestion, CaCl₂ was added to a final concentration of 1.5 mM, taking care to return each sample to 0 °C immediately afterwards. Digestion was performed at 0 °C for 1 h, before being stopped by adding 150 µl of 2X-STOP solution (200 mM NaCl, 20 mM EDTA, 5 mM EGTA, 0.1% NP-40, 40 µg ml⁻¹ glycogen). RNase A was added to a final concentration of 50 µg ml⁻¹ and samples were incubated at 37 °C for 20 min. SDS and proteinase K were then added to final concentrations of 0.1% and 300 µg ml⁻¹, respectively, and samples were incubated at 56 °C for 2 h followed by 68 °C for 16 h to reverse cross-linking. Total DNA was extracted using phenol/chloroform followed by two rounds of ethanol precipitation and DNA size selection (using 0.55× volume of Ampure XP beads relative to the DNA sample volume) to remove the large predominating undigested DNA fragments. Each time, beads were discarded and the supernatant (containing the selected small fragments resulting from MNase digestion) was precipitated with ethanol. After elution in 50 µl TE buffer, samples were quantified and analysed using Qubit and TapeStation assays. CUT&RUN libraries were prepared from 50 ng DNA per sample, using the Accel-NGS 2S Plus DNA Library Kit (Swift) according to the manufacturer's protocol. Paired-end 100-nt sequencing was performed on a HiSeq2500 (Illumina).

Bioinformatics analyses

All data were mapped to the mouse genome mm10, using the BL6-Eij/CAST SNPs from the mouse genome project (v.5 SNP142), and the gene annotation from ensembl (v.92). Analyses were performed in R (v.3.4.2) and Bioconductor (v.3.6). See ref. ²⁴ for more details.

RNA-seq analysis

Reads were trimmed using Trimalore (v.0.4.4), mapped using STAR (2.5.3a, parameters: –outFilterMultimapNmax 1 –outFilterMismatchNmax 999 –outFilterMismatchNoverLmax 0.06 –alignIntronMax 500000 –alignMatesGapMax 500000 –alignEndsType EndToEnd –outSAMattributes NH HI NM MD), and removed when mapping to the mitochondrial genome. Remaining reads were split by allele using SNPsplit (v.0.3.2). Allele-specific and the unassigned bam files were sorted, duplicates removed using picard (v.2.18.2, parameters: REMOVE_DUPLICATES = true ASSUME_SORTED = true) and pooled as the total reads.

Article

Quantification of expression was performed using featureCount (parameters: -p -t exon -g gene_id, -s 1 for stranded RNA-seq of in vitro cell, -s 0 for non-stranded RNA-seq of single embryo). Data were then analysed in R using DESeq2 (v.1.18.1), calculating the sizeFactor on the count of total reads and applying it to the allele-specific counts.

For all RNA-seq analysis (SPEN-degron mouse ES cells, NPCs, and *Spem*-knockout embryos), genes showing less than 10 total allelic reads in at least one sample were discarded from the analysis. Allelic ratios were then computed for genes as follows: $\text{allelic_ratio} = \text{reads}^{\text{B6}} / (\text{reads}^{\text{B6}} + \text{reads}^{\text{Cast}})$. Allelic ratios were then averaged between biological replicates.

RNA-seq analysis of SPEN-degron mouse ES cells. In order to define differential dependencies on SPEN for gene silencing during XCI in mouse ES cells, we removed skewed genes (that is, genes showing allelic ratios outside of a [0.15;0.85] interval in control conditions) from the analysis in Fig. 1d. We then defined a silencing index, translating how much a gene is silenced after 24 h of *Xist* induction with respect to the control condition: $\text{silencing_index} = 1 - (\text{allelic_ratio}_{\text{DOX}} / \text{allelic_ratio}_{\text{control}})$. We next filtered out genes showing less than 10% silencing (that is, $\text{silencing_index} \leq 0.1$) in SPEN non-depleted conditions.

k-means with three clusters was then performed on the raw allelic ratios across control, 24 h DOX and 24 h DOX + auxin conditions. Clustering identified three groups of genes differing by their response to loss of SPEN during *Xist* induction. To define how dependent on SPEN a gene is for silencing, we expressed the silencing defect observed upon loss of SPEN as a fraction of the total silencing that normally occurs in the presence of SPEN. Computationally, this translates in: $\text{Spem_dependence_index} = 1 - (\text{silencing_index}_{\text{DOX+aux}} / \text{silencing_index}_{\text{DOX}})$. A *Spoc*_dependence_index was derived identically.

For integration with *Hdac3*-knockout RNA-seq during XCI, we integrated the SPEN-degron dataset with an *Hdac3*-knockout RNA-seq dataset²⁴ generated from the same mouse ES cell background (TX1072) and at the same time point of *Xist* induction. The dataset was processed identically, and an *Hdac3*-dependence index was also computed as follows: $\text{Hdac3_dependence_index} = 1 - (\text{silencing_index}_{\text{Hdac3KO}} / \text{silencing_index}_{\text{WT}})$.

Spem-knockout E3.5 embryo RNA-seq analysis. Integration of *Spem*-knockout and *Xist*-knockout embryo datasets was performed by integrating our *Spem*-knockout E3.5 female embryo RNA-seq dataset with a *Xist*-knockout single-cell RNA-seq (processed as pseudo-bulk for our analysis) dataset from E3.5 female embryos¹⁰, also generated from a *M. musculus domesticus* × *M. musculus castaneus* mouse background.

SPEN-degron NPC RNA-seq analysis. In NPCs, X-linked genes were defined as escapees if their transcript allelic ratio was greater than 0.15 in at least one condition (0 h, 24 h or 48 h of SPEN depletion).

CUT&RUN bioinformatics analysis

Reads were trimmed using Trimgalore (v.0.4.4), mapped using STAR (2.5.3a, parameters: -outFilterMultimapNmax 1 -outFilterMismatchNmax 999 -outFilterMismatchNoverLmax 0.06 -alignIntronMax 1 -alignMatesGapMax 2000 -alignEndsType EndToEnd -outSAMattributes NH HI NM MD), and removed when mapping to the mitochondrial genome. Remaining reads were split by allele using SNPSplit (v.0.3.2). Allele-specific and the unassigned bam files were sorted, duplicates removed using picard (v.2.18.2, parameters: REMOVE_DUPLICATES = true ASSUME_SORTED = true) and pooled as the total reads. BigWig of coverage files were performed using DeepTools bamCoverage (parameters: -extendReads -binSize 1, with -extendReads 200 for single end data). A scaling factor was calculated as $10^6 / \text{total number of reads}$, and the same factor was given as the parameter -scaleFactor for both allelic signals. Peak calling was performed using macs2 (v.2.1.2.1, parameters for CUT&RUN: -bw 300 -f BAMPE -q 0.01 -keep-dup

auto -broad for CUT&RUN and pol2SS ChIPseq; -bw 300 -f BAMPE -q 0.01 -keep-dup auto -call_summits for other ChIPseq). For quantification of signal in peaks, reads were counted using the featureCounts function from Subread (v.1.28.1, parameters: -p -s 0). Data scaling was performed in R using DESeq2 (v.1.18.1), calculating the sizeFactor on the counts of total reads in 10-kb windows and applying it to the allele-specific counts in peaks.

Peak filtering. SPEN-specific peaks were defined as having $\log_2\text{Fold-Change} \geq 1$ compared to auxin treatment (negative control, SPEN-degraded), and an adjusted *P* value ≤ 0.001 .

Total SPEN enrichment in promoter window. To compare SPEN accumulation among promoters of all X-linked genes in an unbiased manner—including genes that fail to have any peak called at their promoters—we performed DESeq analysis on counts spanning total promoter windows.

Genomic features and integration with RNA-seq. Promoters were defined as ± 2 -kb windows centred around the transcription start sites of genes. Putative active enhancers and their deacetylation kinetics during XCI were obtained from ref.²⁴. Gene-silencing efficiency was determined according to the silencing_index defined in the section 'RNA-seq analysis'. We observed that our silencing_index ranges between 0 and 0.9. Hence, we split this interval in three to define high, medium and low gene-silencing efficiency groups with silencing_index comprising [0.6,0.9], [0.3,0.6] and [0,0.3] respectively.

Integration with publicly available ChIP-seq data. SPEN peaks were intersected with other peaks called from publicly available ChIP-seq data for HDAC3²⁴ (same cellular background, TX1072, 2i + LIF condition), RNAPII-pSS²⁶, CHD4²⁶ and MBD3²⁶ (all in the 2i + LIF condition).

Hi-C analysis

Data were processed with HiC-Pro (v.2.11.0) in allele-specific mode. Only pairs with both reads having MAPQ > 30 were kept. Matrices were made using cooler cload (v.0.8.5) at 1-kb or 10-kb resolution, using HiGlass for visualization and snapshots. Topologically associating domains were called using HicExplorer HicFindTads (parameters: -correctFor-MultipleTesting fdr -minDepth 250000 -maxDepth 4000000 -step 50000 -thresholdComparisons 0.1 -delta 0). The expected value for the Hi-C signal was calculated on the non-allele-specific signal using cooltool compute-expected. Average scaled matrices of observed/expected values for allele-specific signal were produced with Coolpup.py (parameters: -local-rescale-rescale_size 299), using non-allele-specific expected values to normalize both alleles to the same expected values. Average heat maps were plotted using plotpup.py. For quantification, the Hi-C signal was averaged over topologically associating domains upper-triangle for each allele-specific matrix (10 kb) using the hicExplorer hicSummarizeScorePerRegion available at <https://github.com/heard-lab/HiCExplorer> (parameters: -summarizeType mean -rmDiag 1).

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

RNA-seq, Hi-C and CUT&RUN data used in this study have been deposited in the Gene Expression Omnibus under accession number GSE131784. Source Data for Figs. 1–3 and Extended Data Figs. 1, 3, 4 are provided with the paper, either in the form of supplementary tables or source data files.

31. de Vries, W. N. et al. Expression of Cre recombinase in mouse oocytes: a means to study maternal effect genes. *Genesis* **26**, 110–112 (2000).
32. Zyllicz, J. J. et al. G9a regulates temporal preimplantation developmental program and lineage segregation in blastocyst. *eLife* **7**, e33361 (2018).
33. Tang, F. et al. RNA-seq analysis to capture the transcriptome landscape of a single cell. *Nat. Protoc.* **5**, 516–535 (2010).
34. Huang, Y. et al. Stella modulates transcriptional and endogenous retrovirus programs during maternal-to-zygotic transition. *eLife* **6**, e22345 (2017).
35. Belaghal, H., Dekker, J. & Gibcus, J. H. Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods* **123**, 56–65 (2017).
36. Chen, J. et al. High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed by single virion analysis. *Proc. Natl Acad. Sci. USA* **106**, 13535–13540 (2009).
37. Barau, J. et al. The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. *Science* **354**, 909–912 (2016).
38. Pouillet, P., Carpentier, S. & Barillot, E. myProMS, a web server for management and validation of mass spectrometry-based proteomic data. *Proteomics* **7**, 2553–2556 (2007).
39. Valot, B., Langella, O., Nano, E. & Zivy, M. MassChroQ: a versatile tool for mass spectrometry quantification. *Proteomics* **11**, 3572–3577 (2011).

Acknowledgements We thank K. Ancelin for support throughout the project and help with in vivo experiments; J. Barau, D. Holoch and R. Margueron for support and help with the project; M. Carrara for critical reading of the manuscript; and members of the Heard laboratory for discussions. We thank E. Nora for sharing the OsTIR1 and AID-targeting plasmids; T. Honjo for sharing the *Spen*^{flax} mouse line; O. Masui for sharing the *Xist*-Bgl stem-loop targeting constructs and L. Lavis for sharing Halo-JF646 and JF549 with us. We also thank the imaging platform (A. Dauphin and PICT-IBISA (UMR3215/U934)) and the protein purification and

sequencing platforms of Institut Curie as well as L. Villacorta, J. Provaznik and V. Benes of GeneCore at EMBL. This work was funded by a Boehringer Ingelheim doctoral fellowship (20017-2019 to F. Dossin), an ERC Advanced Investigator award (ERC-ADG-2014 671027 to E.H.), Labellisation La Ligue (to E.H.), ANR (DoseX 2017: ANR-17-CE12-0029, Labex DEEP: ANR-11-LBX-0044, ABS4NGS: ANR-11-BINF-0001, and part of the IDEX PSL: ANR-10-IDEX-0001-02 PSL to E.H.), a Sir Henry Wellcome Postdoctoral Fellowship (201369/Z/16/Z to J.J.Z.), 'Région Ile-de-France' and Fondation pour la Recherche Médicale grants (to D.L.), EMBO long-term fellowships (ALTF 549-2014 to I.P., ALTF 301-2015 to T.C.), and Fondation pour la Recherche Médicale (SPF 20140129387 to I.P.) and NIH (HG003143 to J.D.) grants. J.D. is an investigator of the Howard Hughes Medical Institute.

Author contributions F. Dossin and E.H. conceived the experiments and F. Dossin performed them unless stated otherwise. T.M. provided the *Spen*-knockout mice and J.J.Z. performed the embryo experiments. E.H. and I.P. supervised the work. J.R. established the *Xist* stem-loop cell line and acquired live-cell images. F. Dossin, J.R. and V.K. analysed live-cell imaging experiments. V.K. wrote the image analysis software. F. Dossin and S.C. processed and analysed the data. A.L.S. cloned the SPEN cDNA truncation constructs. T.C. helped with some experiments and M.A. derived NPC clones. Y.Z. prepared Hi-C libraries in the laboratory of J.D. F. Dingli carried out the mass spectrometry experimental work and D.L. supervised mass spectrometry experiments and data analysis. F. Dossin and E.H. wrote the manuscript with input from all authors.

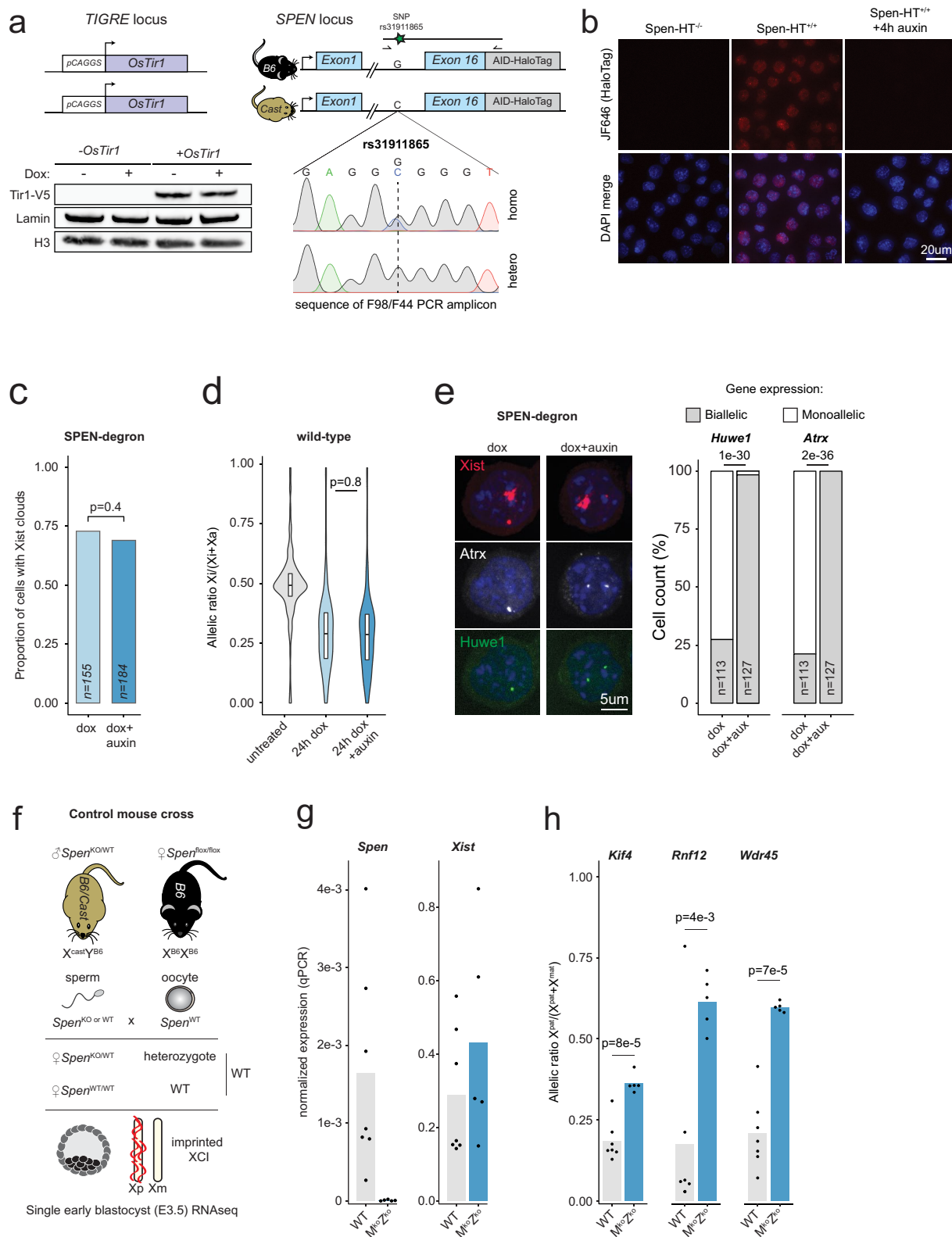
Competing interests The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-1974-9>.

Correspondence and requests for materials should be addressed to E.H.

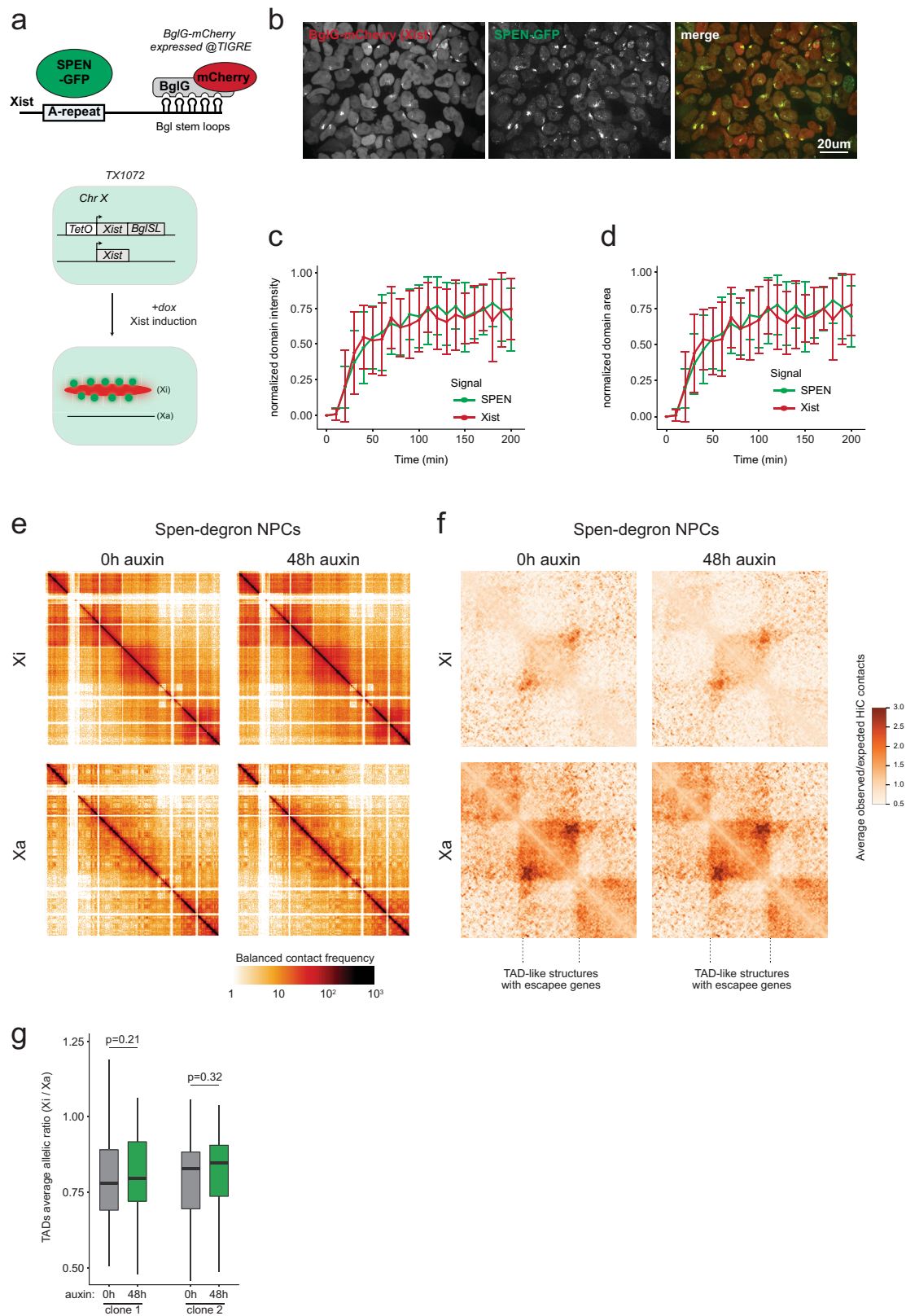
Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | SPEN mediates gene silencing across the entire X chromosome in vitro and in vivo. **a**, Schematic representation of the SPEN-degron genotype with AID-HaloTag insertions in frame with the C terminus of endogenous SPEN. Targeted homozygous insertion of V5-tagged OsTIR1 at the *TIGRE* locus (top left) results in its constitutive protein expression as assessed by western blot (bottom left). Right, Sanger sequencing results for a PCR amplicon specific to AID-HaloTag insertions and covering a SNP outside of the recombined left homology arm. Detection of both alleles in the amplicon confirms homozygous AID knock-in. **b**, Fixed-cell imaging of HaloTag in wild-type cells (left), in SPEN-degron mouse ES cells (middle) and in SPEN-degron mouse ES cells exposed to auxin for 4 h (right). Cells were labelled with Halo-JF646 before fixation. SPEN-Halo is properly localized to the nucleus, and is depleted upon auxin treatment. This experiment was repeated at least twice with similar results. **c**, Bar graph showing the proportion of cells displaying *Xist* RNA clouds (quantified using RNA FISH) before and after degradation of SPEN (n , number of cells counted; χ^2 test). **d**, Violin plot showing the distribution of

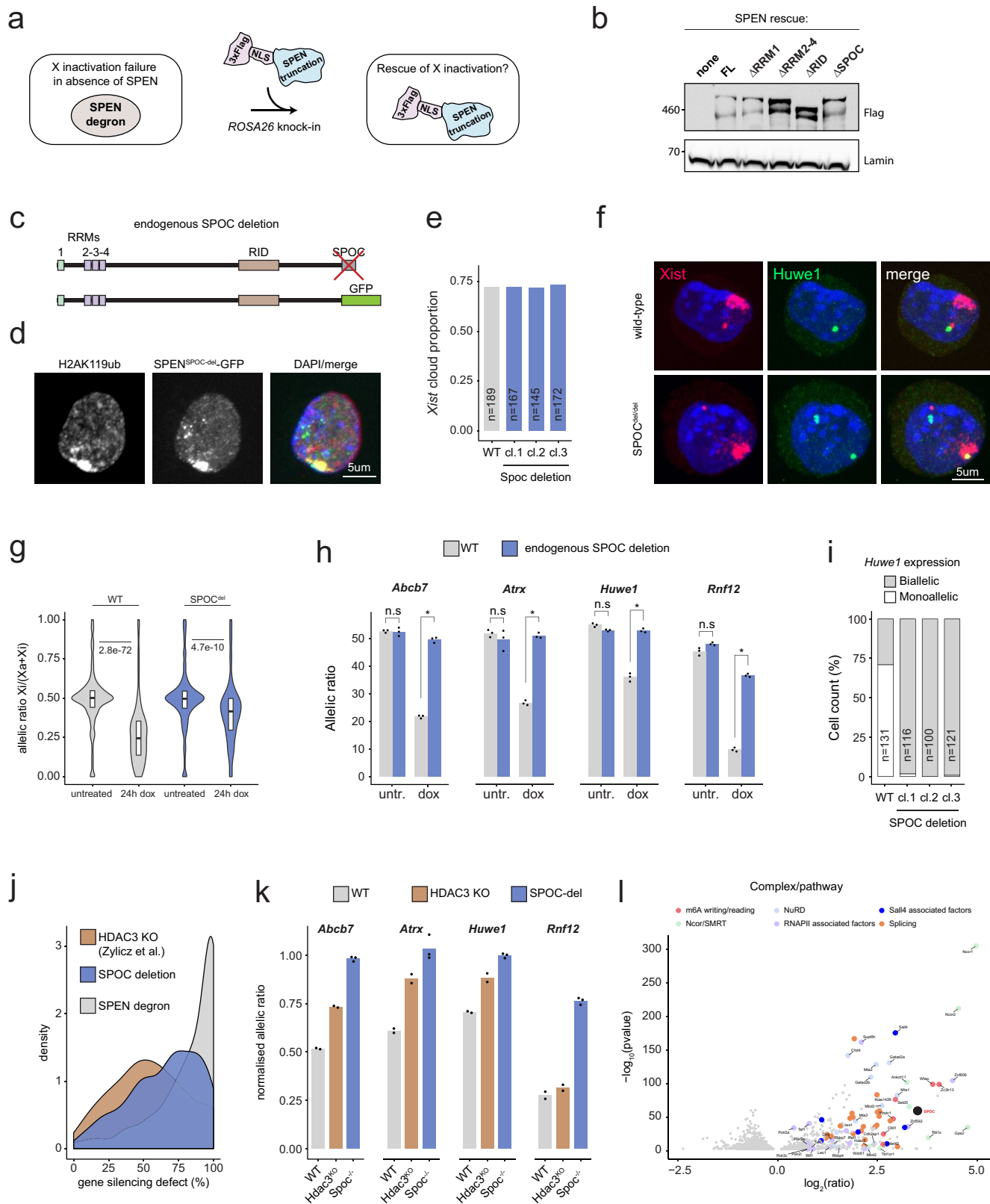
X-chromosomal transcript allelic ratios (obtained by RNA-seq) after 0 h DOX, 24 h DOX or 24 h DOX + auxin treatment in wild-type SPEN-degron mouse ES cells. Horizontal lines denote the median, box limits correspond to upper and lower quartiles, averages of two independent clones shown, $n = 434$ genes, two-sided Student's t -test. **e**, RNA FISH experiments for *Xist* (red) and two X-linked genes: *Atrx* (grey) and *Huwe1* (green), in SPEN-degron mouse ES cells treated with DOX only, or DOX in combination with auxin for 24 h. The proportion of *Atrx/Huwe1* monoallelic and biallelic expression among *Xist*-expressing cells is shown (n , number of cells counted; χ^2 test). **f**, Illustration of the control hybrid mouse crossbreeding scheme for the experiment shown in Fig. 1g, h. **g**, Quantitative PCR (qPCR) analysis of *Spn* and *Xist* transcripts in wild-type ($n = 7$) and maternal-zygotic *Spn*-knockout ($n = 5$) E3.5 embryos. **h**, Pyrosequencing assay of three X-linked transcripts in maternal-zygotic *Spn*-knockout ($n = 5$) and wild-type ($n = 7$) E3.5 embryos (two-sided Student's t -test). In **g**, **h**, bars show the mean value and individual data points are shown as dots.



Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | SPEN localizes to the X chromosome immediately upon *Xist* upregulation and throughout the stages of XCI, but is dispensable for maintenance of X-linked gene silencing. **a**, Scheme of the strategy for live-cell imaging of SPEN protein and *Xist* RNA. **b**, Live-cell snapshot after 16 h of *Xist* induction in the cell line shown in **a**. This experiment was repeated at least twice with similar results. **c, d**, Kinetics of total intensity (**c**) and area (**d**) of *Xist* (red) and SPEN (green) domains over time during *Xist* induction. The data in **c, d** are the averages of 27 tracked cells. Error bars indicate standard deviation. Images were acquired every 10 min. Time point 1 is defined as the earliest time at which a SPEN or *Xist* domain is detected in each cell. Intensity and area values were respectively normalized to the maximum value reached for each signal

(SPEN and *Xist*). **e**, Hi-C map of the inactive (top) and active (bottom) X chromosomes (resolution, 1.024 Mb) in NPCs after 0 h or 48 h of auxin-mediated SPEN depletion. **f**, Heat map of the average contact enrichment on scaled topologically associating domains containing escapees in NPCs after 0 h or 48 h of auxin-mediated SPEN depletion. **g**, Quantification of the allelic ratio (inactive/active X chromosome) of the Hi-C signal within topologically associating domains ($n = 37$) shown in **f**, after 0 h or 48 h of auxin-mediated SPEN depletion. Horizontal lines denote the median, box limits correspond to upper and lower quartiles, two-sided Wilcoxon rank-sum test. In **e, f**, averages of two independent clones are shown.

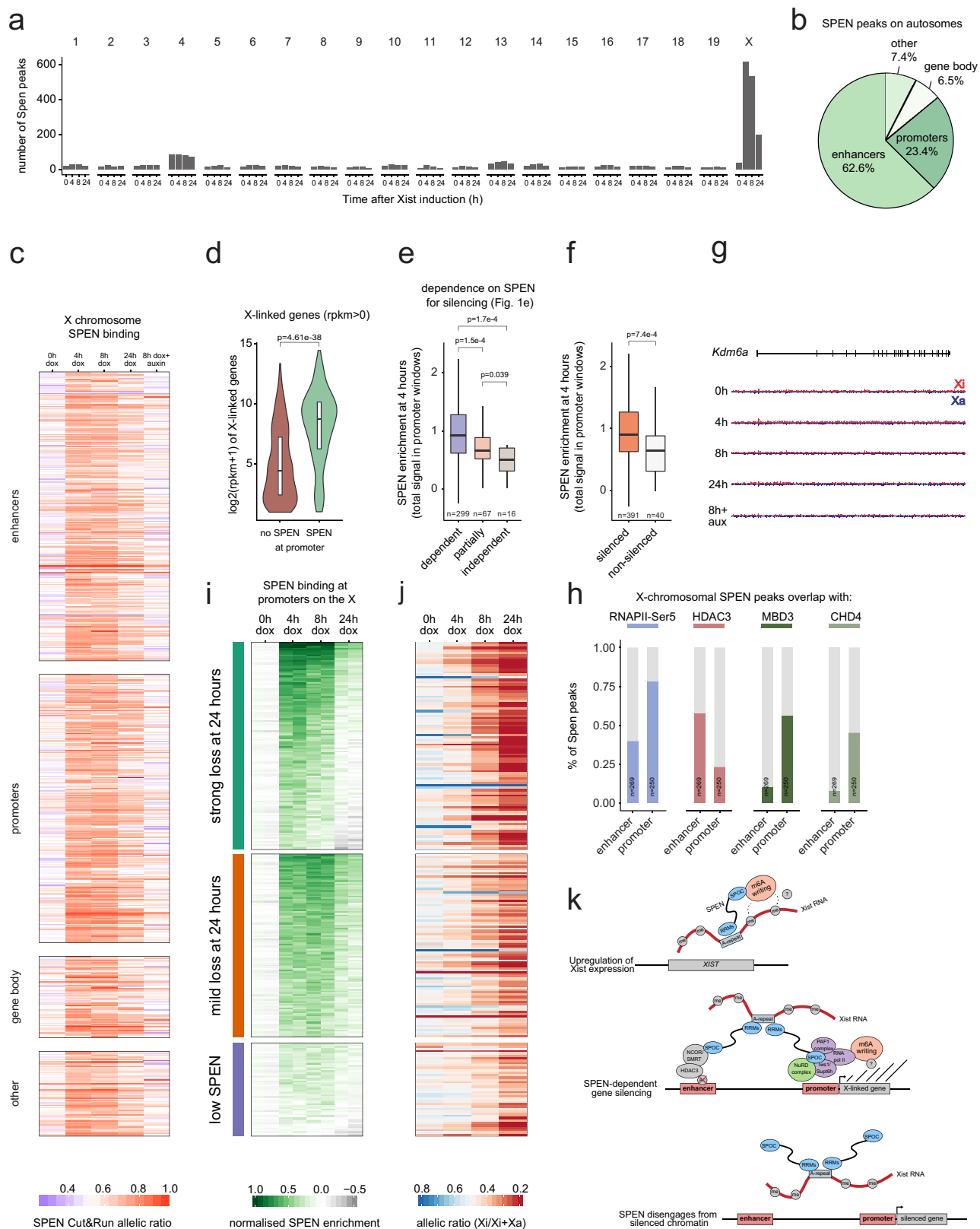


Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | The SPOC domain of SPEN mediates gene silencing

and interacts with multiple molecular pathways. **a**, Scheme of complementation strategy. **b**, Western blot detection of overexpressed 3×Flag-tagged SPEN protein rescue fragments. **c**, Scheme showing endogenous deletion of SPOC. **d**, Sub-nuclear localization of endogenous SPEN lacking its SPOC domain upon *Xist* RNA induction. The inactive X chromosome is identified using immunofluorescence detection of H2AK119ub1. **e**, Bar graph showing the proportion of cells with *Xist* RNA clouds (assayed by RNA FISH) in wild-type cells and three independent SPOC-deletion clones after induction of *Xist* for 24 h (*n*, number of counted cells). **f**, RNA FISH for *Xist* (red) and *Huwe1* (green) in SPOC-deletion and wild-type cells treated with DOX for 24 h. **g**, Violin plot showing the distribution of X-chromosomal transcript allelic ratios (measured by RNA-seq) after 0 h or 24 h DOX treatment in wild-type and SPOC-deletion mouse ES cells. Horizontal lines denote the median, box limits correspond to upper and lower quartiles, averages of three independent clones shown, *n* = 469 genes, two-sided Student's *t*-test. **h**, Bar graph of transcript allelic ratios (obtained from pyrosequencing) for four X-linked genes in SPOC-deletion (blue) or wild-type (grey) cells. Bars show mean values for three

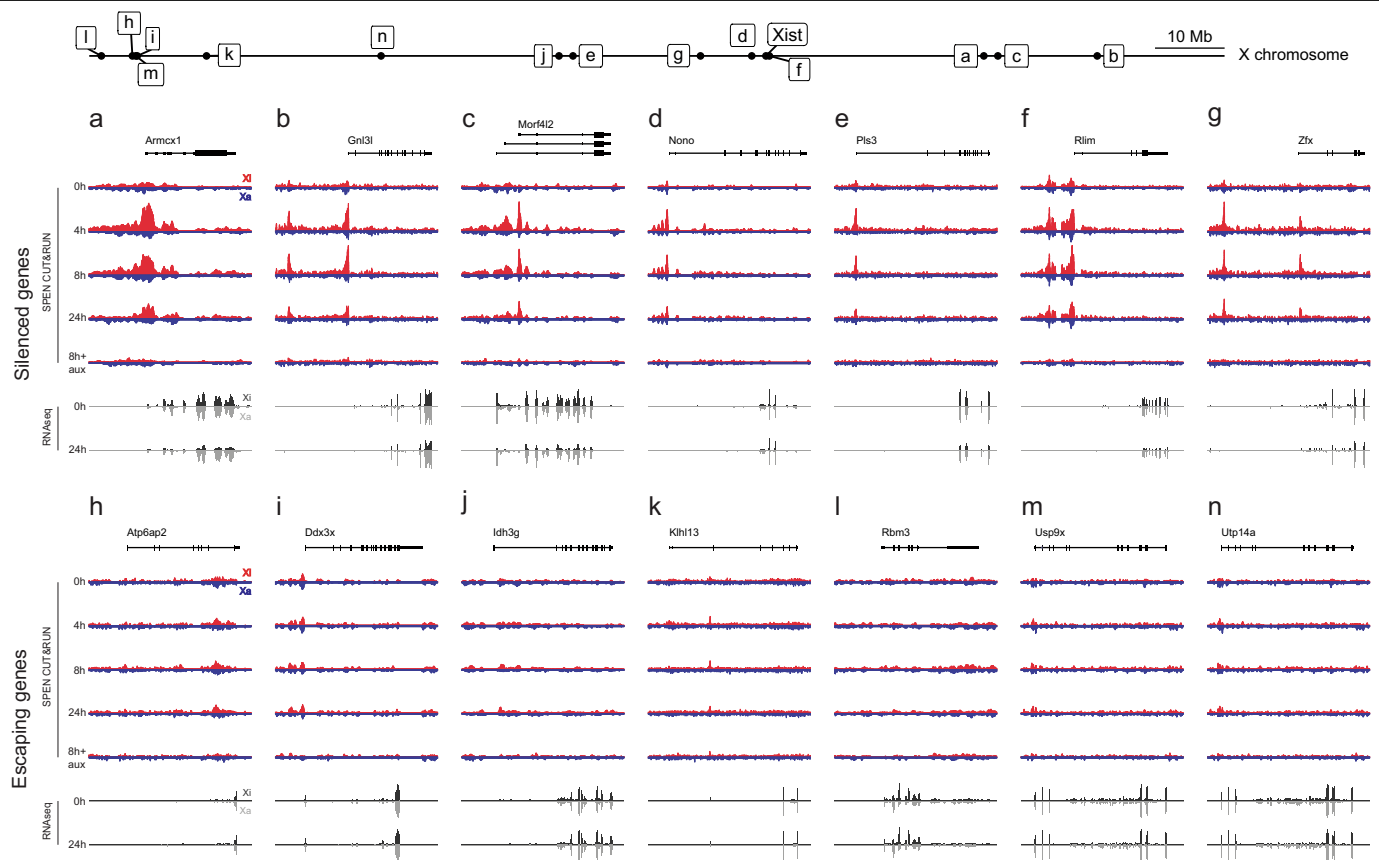
independent SPOC-deletion clones ($^*P < 10^{-4}$, two-sided Student's *t*-test). **i**, Bar graph showing the proportion of cells expressing *Huwe1* monoallelically (white) or biallelically (grey), assayed by RNA FISH, in wild-type cells and in three independent SPOC-deletion clones after induction of *Xist* for 24 h (*n*, number of counted cells). **j**, Density plot showing the distribution of gene silencing defects (see Methods) observed across the X chromosome in RNA-seq data from HDAC3-knockout²⁴ SPEN-degron and SPOC-deletion (this study) ES cells after 24 h of *Xist* induction. **k**, Bar graph of normalized allelic ratios (obtained from pyrosequencing) for four X-linked genes in HDAC3-knockout (brown), SPOC-deletion (blue) and wild-type (grey) cells after 24 h of *Xist* induction. Bars show mean values for two independent HDAC3 clones and three independent SPOC deletion clones; individual data points are shown. **l**, Volcano plot of fold changes in GFP-pull-down (BglG–GFP–SPOC compared with BglG–GFP) and their adjusted *P* values (Benjamini–Hochberg procedure, see Methods for statistical analysis). Quantitative label-free mass spectrometric analysis was performed on four independent biological replicates. In **b**, **d**, **f**, experiments were repeated at least twice with similar results.



Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | SPEN is recruited by *Xist* to active gene promoters and enhancers where it silences transcription and subsequently disengages from chromatin. **a**, Bar graph showing the number of SPEN peaks on each chromosome after 0 h, 4 h, 8 h and 24 h of *Xist* induction in mouse ES cells. **b**, Annotation of SPEN peaks on autosomes. **c**, Heat map showing allelic ratios at SPEN peaks during XCI among different X-linked genomic features. **d**, Violin plot showing expression (RPKM) of genes accumulating SPEN ($n = 259$) or not accumulating SPEN ($n = 689$) at their promoters. Genes showing 0 RPKM were excluded from this plot. **e**, Box plots showing SPEN enrichment after 4 h of *Xist* induction within promoter windows of genes grouped on the basis of their level of dependency on SPEN for gene silencing (see Fig. 1e). **f**, Box plots showing SPEN enrichment after 4 h of *Xist* induction within promoter windows of genes grouped on the basis of whether or not they are silenced at 24 h of *Xist*

induction (see Methods). In **d–f**, data were analysed using the two-sided Wilcoxon rank-sum test, horizontal lines denote the median, box limits correspond to upper and lower quartiles. **g**, UCSC Genome Browser allele-specific track showing SPEN binding around *Kdm6a*, an escaping gene (blue, Cast-Xa; red, B6-Xi; all tracks are scaled identically). **h**, Bar graphs showing overlap between SPEN-binding sites and the binding sites of four different factors at X-linked enhancers and promoters. **i, j**, Heat maps showing normalized SPEN enrichment (\log_2) at promoters (both replicates are shown) (**i**) and gene silencing kinetics (allelic ratio) during XCI (**j**) within three groups of X-linked genes showing different dynamics of SPEN accumulation and loss. **k**, Schematic of the function of SPEN in XCI. In **a–f, h–j**, data are from two biological replicates.



Extended Data Fig. 5 | UCSC Genome Browser allelic tracks of SPEN binding and transcript expression at X-linked genes. a–n, Top, Genome Browser allelic tracks of SPEN binding (from CUT&RUN) at silenced genes (a–g) and non-silenced genes (h–n) during a time course of *Xist* induction in mouse ES cells (blue, Cast-Xa; red, B6-Xi; scaled identically within each panel). Bottom, allelic

tracks of transcript expression (from RNA-seq) at 0 h and 24 h of *Xist* induction in mouse ES cells (light grey, Cast-Xa; black, B6-Xi; scaled identically within each panel). The relative position of each gene along the X chromosome is shown at the top of the figure.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- | | | |
|-------------------------------------|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | A description of all covariates tested |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

All microscopy images were acquired either with an Inverted Confocal Spinning Disk Roper/Nikon for fixed cells or a Super-resolution microscope OMX (Applied Precision Incorporation, DeltaVision) for live cells. Sequencing data was collected using the Illumina platform. Pyrosequencing data was collected using the PyroMark Q24 System from Qiagen. QPCR was performed on a ViiA 7 Real-Time PCR System from Thermo Fisher Scientific. Mass Spectrometry was performed on Thermo Scientific Orbitrap Fusion Tribrid MS from Thermo Fisher Scientific. Western blot (chemiluminescent) images were collected using a ChemiDoc MP from BioRad.

Data analysis

Trimalore (v 0.4.4), STAR (2.5.3a), SNPsplit (v 0.3.2), picard (v2.18.2), DESeq2 (v1.18.1), R (3.4.1), featureCount (1.6.3), Fiji (2.0.0), dplyr (0.7.4), readr (1.1.1), tidyr (0.8.0), ggplot2 (2.2.1), macs2 (v 2-2.1.2.1), Subread (v1.28.1), HiC-Pro (v2.11.0)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data generated in this study are available on GEO database under the number GSE131784.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was not predetermined. For mouse experiments, we used a sample size commonly used and accepted for basic statistical inference while using an justifiable number of mice. For tissue culture based experiments, at least two independent clones were systematically assessed for each genotype. Typically, for the SPOC tethering experiment, 4 independent clones were characterized with very little variation being observed between them.
Data exclusions	No data were excluded from the analysis, with the exception of live imaging analysis, from which dying cells, as well as cells which moved outside from the field of view during movie acquisition were removed.
Replication	All attempts at replication were successful and noted in the relevant figure legend.
Randomization	Samples were not randomized, given that samples were grouped according to their respective genotypes.
Blinding	For most experiments, no blinding was performed as most measurements were derived from third party machines/software, and hence not affected by subjective interpretation. For RNA FISH image analysis (counting of Xist RNA clouds and X-linked gene pinpoints), each experimental condition was blinded from the first author F. Dossin during counting.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used	Epitope, Antibody, reference, Application, Dilution HaloTag, Promega cat. #G9211, WB, 1/1000 V5, Sigma cat. #V8012, WB, 1/2500, Flag, Sigma cat. #F1804, WB and IF, 1/1000 and 1/200 GFP, Abcam cat. #ab290 LotGR3222604-1, Cut&Run and IF, 1/200 GFP, Roche cat. #11814460001, WB, 1/1000 H2Ak119ub1, Cell Signaling cat. #8240, IF, 1/500 Lamin B1, Abcam cat. #ab16048, WB, 1/3000 PCNA, DAKO cat. #M0879, WB, 1/3000 Ncor1, Abcam cat. #ab2482 LotGR320472-6, WB, 1/1000 Ncor2, Abcam cat. #ab5802 LotGR259451-18, WB, 1/1000 Mta1, Cell Signaling cat. #5647, WB, 1/1000 Wtap, Proteintech cat. #10200-1-AP, WB, 1/1000 Mettl3, Abcam cat. #ab195352, WB, 1/1000 Hdac3, SantaCruz cat. #sc-376957, WB, 1/1000 Rpb1, Abcam cat. #ab817, WB, 1/2000
Validation	The HaloTag antibody is validated for western blot in Fig. 1b, with specific signal disappearing upon auxin treatment, and reappearing upon removal of auxin from the culture medium. The V5 antibody is validated for Western blot in Extended Data Fig. 1a, with specific signal being observed only in cells expressing

V5-tagged Tir1.

The Flag antibody is validated for immunofluorescence in Fig. 3c, with specific nuclear signal being observed only in cells expressing Flag-tagged SPEN truncations.

The GFP antibody is validated for CUT&RUN by showing specific genomic signal for GFP-tagged SPEN, which disappears upon treatment of cells with auxin.

The H2AK119ub1 has been validated for immunofluorescence in Zylitz et al., 2019

The Lamin B1 antibody have been KO validated by abcam.

The PCNA antibody is cited 361 times according to CiteAb.

The Ncor1/Ncor2 antibodies are validated for western blot in Fig3. h, as showing increased signal upon immunoprecipitation of SPOC.

The Mta1 antibody has been cited 12 times previously according to CiteAb.

The Wtap antibody have been KO validated by ProteinTech.

The Mettl3 antibody have been KO validated by Abcam.

The Hdac3 antibody has been cited 9 times according to CiteAb.

The Rpb1 antibody has been cited 280 times previously according to CiteAb.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

All cell lines were derived from the TX1072 female mouse embryonic stem cell line (Schulz et al., 2014)

Authentication

None of the cell lines were authenticated

Mycoplasma contamination

All cell lines tested negative for mycoplasma contamination

Commonly misidentified lines
(See [ICLAC](#) register)

cells used are not in the ICLAC database

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

We used adult animals (from 6 weeks to 3 months for females, and 7 weeks to 1 year for males) for producing preimplantation embryos (first 4 days post mating). For Spen matings a published conditional allele was used (Yabe et al., 2007). For oocyte deletions published Rosa26:Zp3-Cre allele was used (DeVries et al., 2000). F1 hybrid Spen+/- males were obtained by crossing Spen+/+ CAST/EiJ females with Spen+/- C57BL/6J males. For Spen maternally deleted embryos, Spenflox/flox Zp3-Cre+ve C57BL/6J females were crossed with Spen+/- F1 hybrid males. For Spen control embryos, Spenflox/flox Zp3-Cre-ve C57BL/6J females were crossed with Spen+/- F1 hybrid males.

Wild animals

this study does not contain any wild animals

Field-collected samples

this study does not contain animals collected from the fields.

Ethics oversight

Animal care and use for this study were performed in accordance with the recommendations of the European community (2010/63/UE). All experimental protocols were approved by the ethics committee of Institut Curie CEEA-IC118 under the number APAFIS#8812-2017020611033784v2 given by national authority in compliance with the international guidelines.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

NEDD8 nucleates a multivalent cullin–RING–UBE2D ubiquitin ligation assembly

<https://doi.org/10.1038/s41586-020-2000-y>

Received: 16 August 2019

Accepted: 9 January 2020

Published online: 12 February 2020



Kheewoong Baek¹, David T. Krist^{1,3}, J. Rajan Prabu¹, Spencer Hill², Maren Klügel¹, Lisa-Marie Neumaier¹, Susanne von Gronau¹, Gary Kleiger² & Brenda A. Schulman¹✉

Eukaryotic cell biology depends on cullin–RING E3 ligase (CRL)-catalysed protein ubiquitylation¹, which is tightly controlled by the modification of cullin with the ubiquitin-like protein NEDD8^{2–6}. However, how CRLs catalyse ubiquitylation, and the basis of NEDD8 activation, remain unknown. Here we report the cryo-electron microscopy structure of a chemically trapped complex that represents the ubiquitylation intermediate, in which the neddylated CRL1^{β-TRCP} promotes the transfer of ubiquitin from the E2 ubiquitin-conjugating enzyme UBE2D to its recruited substrate, phosphorylated IκBα. NEDD8 acts as a nexus that binds disparate cullin elements and the RING-activated ubiquitin-linked UBE2D. Local structural remodelling of NEDD8 and large-scale movements of CRL domains converge to juxtapose the substrate and the ubiquitylation active site. These findings explain how a distinctive ubiquitin-like protein alters the functions of its targets, and show how numerous NEDD8-dependent interprotein interactions and conformational changes synergistically configure a catalytic CRL architecture that is both robust, to enable rapid ubiquitylation of the substrate, and fragile, to enable the subsequent functions of cullin–RING proteins.

CRLs orchestrate numerous eukaryotic processes—including transcription, signalling, cell division and differentiation—and CRL dysregulation underlies many pathologies. The activities of these enzymes depend on coordinated but dynamic interactions between dedicated cullin–RING complexes and several regulatory partner proteins¹. Cullins (CULs) 1–5 bind cognate RING-containing partners (RBX1 or RBX2) through a conserved intermolecular cullin/RBX (hereafter C/R) domain, in which a CUL β-sheet stably embeds an RBX strand⁷. On one side of the C/R domain, the CUL N-terminal domain can associate interchangeably with numerous substrate-recruiting receptors. As examples, human CUL1–RBX1 binds around 70 SKP1-F-box protein complexes and CUL4–RBX1 binds around 30 DDB1–DCAF complexes, forming the E3 enzymes CRL1^{F-box protein} or CRL4^{DCAF}, respectively (in which F-box protein and DCAF represent substrate receptors for a given CRL)^{8–11}. The C-terminal WHB domain of the CUL protein and the RING domain of the RBX protein emanate from the other side of the C/R domain. To achieve E3 ligase activity, the RING domain recruits one of several ubiquitin-carrying enzymes, which presumably use distinct mechanisms to transfer ubiquitin to receptor-bound substrates.

CRLs are regulated by reversible NEDD8 modification of a specific lysine residue within the WHB domain of CULs. Although it has approximately 60% sequence identity to ubiquitin, NEDD8 uniquely activates CRL-dependent ubiquitylation². NEDD8 has been suggested to have multiple roles in catalysis—including assisting in the recruitment of ubiquitin-carrying enzymes, facilitating juxtaposition of the substrate and ubiquitylation active site, and promoting conformational changes—although the structural mechanisms of these effects

remain unknown^{3–5,12}. NEDD8 also stabilizes cellular CRLs by blocking the exchange factor CAND1 from ejecting substrate receptors from unneddyated CUL–RBX complexes¹³. Neddylation controls around 20% of ubiquitin-mediated proteolysis and presumably many nondegradative functions of ubiquitin, and an inhibitor of NEDD8 (MLN4924, also known as Pevonedistat) blocks HIV infectivity and is in clinical trials as an anticancer agent^{6,14}.

Here we determine structural mechanisms that underlie ubiquitylation by human neddyated CRL1^{β-TRCP} and E2 enzymes from the UBE2D family, in which the F-box protein β-TRCP recruits a specific phosphodegron motif in substrates including β-catenin and IκBα^{15–20}. UBE2D knockdown stabilizes the CRL1^{β-TRCP} substrate IκBα²¹, whereas mutations that impair the ubiquitylation of β-catenin by CRL1^{β-TRCP} promote tumorigenesis²². Furthermore, hijacking of CRL1^{β-TRCP} enables HIV to evade host immunity²³, and deamidation of Gln40 of NEDD8 by an enteropathogenic and enterohaemorrhagic *Escherichia coli* effector results in the accumulation of the CRL1^{β-TRCP} substrate IκBα as well as substrates of other CRLs^{24–26}.

NEDD8 activation of ubiquitylation

We used rapid quench-flow methods to obtain kinetic parameters for CRL1^{β-TRCP} and UBE2D catalysed ubiquitylation of model substrates (phosphopeptides from β-catenin and IκBα, containing single acceptor lysines). We found that NEDD8 substantially stimulates the reaction, by nearly 2,000-fold (Fig. 1a, Extended Data Fig. 1, Extended Data Table 1). Performing experiments under conditions that allow

¹Department of Molecular Machines and Signaling, Max Planck Institute of Biochemistry, Martinsried, Germany. ²Department of Chemistry and Biochemistry, University of Nevada, Las Vegas, Las Vegas, NV, USA. ³Present address: Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA. ✉e-mail: schulman@biochem.mpg.de

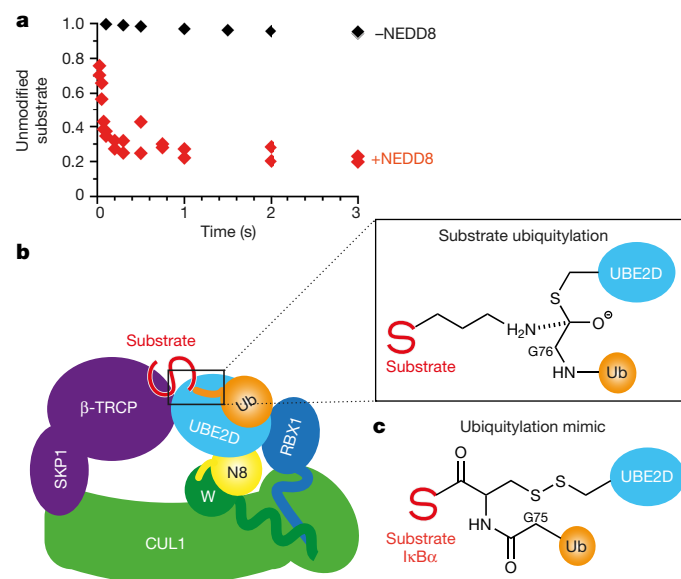


Fig. 1 | Role of NEDD8 and strategy for the visualization of dynamic ubiquitin transfer from UBE2D to substrate by neddylation of CUL1. **a**, Effect of the neddylation of CUL1 on CRL1 ^{β -TRCP}-catalysed ubiquitin transfer from UBE2D to a radiolabelled IkB α -derived peptide substrate. The plots show the proportion of substrate remaining during pre-steady-state rapid quench-flow ubiquitylation reactions with saturating UBE2D3 and either unneddylated or neddylation of CRL1 ^{β -TRCP}. The symbols show the data from independent experiments ($n = 2$ technical replicates). **b**, Schematic representing substrate priming by neddylation of CRL1 ^{β -TRCP} and UBE2D-Ub. The inset shows the transition state during ubiquitylation. **c**, Chemical mimic of the ubiquitylation intermediate, in which surrogates for the active site of UBE2D, the C terminus of ubiquitin and the ubiquitin acceptor site on the IkB α -derived substrate peptide are simultaneously linked.

for multiple UBE2D turnover events enabled us to quantify the effects of neddylation on substrate ‘priming’, in which ubiquitin is ligated directly to the substrate, compared with ‘chain elongation’, in which it is linked to a substrate-linked ubiquitin. The individual rates for the linkage of successive ubiquitins during polyubiquitylation showed that NEDD8 activates both substrate-priming and chain-elongation reactions. However, ubiquitin ligation to a substrate is tenfold faster than ubiquitin ligation to a substrate-linked ubiquitin, suggesting that neddylation of CRL1 ^{β -TRCP}—together with UBE2D—optimally catalyses substrate-priming reactions (Extended Data Table 1).

Cryo-EM reveals cullin–RING dynamics

Ubiquitin transfer from a RING-docked UBE2D-Ub intermediate (in which - indicates a thioester bond or thioester-bond mimic, Ub indicates ubiquitin) to a substrate that is bound to an F-box protein was difficult to rationalize from previous structural models^{3,7,20,27} and from our cryo-electron microscopy (cryo-EM) reconstructions of unneddylated and neddylation substrate-bound CRL1 ^{β -TRCP} (Extended Data Fig. 2, Extended Data Table 2). We observed a well-resolved ‘substrate-scaffolding module’ that resembles models based on crystal structures of substrate-bound SKP1– β -TRCP and the portion of an F-box–SKP1–CUL1–RBX1 complex that includes the N-terminal domain of CUL1 and the intermolecular C/R domain^{7,20}. However, for the RING domain of RBX1 and the WHB domain of CUL1—with or without covalently linked NEDD8—density is either lacking or visualized only at low contour, in varying positions in different classes. These domains apparently sample multiple orientations, and it is therefore difficult to conceptualize rapid ubiquitylation of a flexible substrate by uncoordinated nanometre-scale motions of RBX1-activated UBE2D-Ub (Extended Data Fig. 2).

Capturing CRL–substrate ubiquitylation

Substrate priming involves fleeting simultaneous linkage of the active site of UBE2D, the C terminus of ubiquitin and the substrate (Fig. 1b). We chemically linked surrogates for these entities to form a stable mimic of the transition state, which avidly binds neddylation of CRL1 ^{β -TRCP} (Fig. 1c, Extended Data Fig. 3a–d). After screening several complexes by cryo-EM, we obtained a reconstruction at 3.7 Å resolution that showed our proxy for a UBE2D-Ub–IkB α substrate intermediate bound to a hyperactive version of neddylation of CRL1 ^{β -TRCP} (Fig. 2a, Extended Data Figs. 3, 4, Extended Data Table 2).

This complex, representing the neddylation of CRL1 ^{β -TRCP}–UBE2D-Ub–substrate intermediate, explains rapid ubiquitylation through unprecedented neddylation of cullin–RING arrangements (Extended Data Fig. 5). NEDD8 is nearly encircled by interactions; it binds the WHB domain of CUL1—to which it is linked—in an ‘activation module’, and positions a ‘catalytic module’ relative to the substrate-scaffolding module to juxtapose the active site and the substrate (Fig. 2b–d).

In the catalytic module, RBX1 binds UBE2D-Ub in the canonical RING-activated ‘closed’ conformation, in which noncovalent interactions between UBE2D and ubiquitin allosterically activate the thioester bond between them^{28–30}. Compared with previously isolated RING–UBE2D-Ub structures^{28–30}, the neddylation of CRL1 ^{β -TRCP}–UBE2D-Ub–substrate intermediate shows additional density corresponding to the substrate proxy along a trajectory to the ubiquitylation active site (Fig. 2c, Extended Data Figs. 6a, 7b). A UBE2D groove seems to engage the substrate polypeptide in a manner poised to assist in projecting adjacent lysines into the active site. This engagement of substrate polypeptide may contribute to the ability of UBE2D to ubiquitylate a broad range of proteins³¹.

The structure revealed that the distance between the β -TRCP-bound phosphodegron of IkB α and the UBE2D-Ub active site is around 22 Å, which is compatible with the spacing between this motif and potential acceptor lysines in many substrates³² (Extended Data Fig. 6). We could therefore make the following predictions: firstly, peptide substrates with sufficient residues (for example, 13 and 9) between the phosphodegron and the acceptor lysine to span a distance of 22 Å should be rapidly primed in a NEDD8-dependent manner; secondly, a peptide substrate with too few spacer residues (for example, 4) to span this gap should be severely impaired for priming by neddylation of CRL1 ^{β -TRCP} and UBE2D, but that the addition of a ubiquitin could satisfy geometric constraints and enable further polyubiquitylation; and thirdly, the substrates should show little difference in UBE2D-mediated priming with unneddylated CRL1 ^{β -TRCP}. Comparing kinetic parameters for peptide substrates with 13-, 9-, and 4-residue spacers confirmed these predictions, with the priming rate of the latter peptide with neddylation of CRL1 ^{β -TRCP} reduced below the limit of our quantification (Extended Data Fig. 6, Extended Data Table 1).

NEDD8 coordinates ubiquitin ligation assembly

NEDD8 is covalently linked to the WHB domain from CUL1, and together they form a globular activation module. A NEDD8 groove, comprising the Ile36/Leu71/Leu73 hydrophobic patch and the C-terminal tail, embraces the hydrophobic face of the isopeptide-bound CUL1 helix (Figs. 2d, 3a, b). At the centre, Gln40 of NEDD8 contacts CUL1, the isopeptide bond, and the C-terminal tail of NEDD8 in a buried polar interaction that is typical of such organizing apolar interfaces³³. This rationalizes how pathogenic bacterial effectors that catalyse Gln40 deamidation impair CRL1-dependent ubiquitylation^{12,24–26}: the resultant negative charge would destroy the CUL1–NEDD8 interface.

The activation module binds the catalytic module, which explains how neddylation helps CRL1 ^{β -TRCP} to recruit UBE2D in cells³⁴. The Ile44 hydrophobic patch of NEDD8 engages the ‘backside’ of UBE2D,

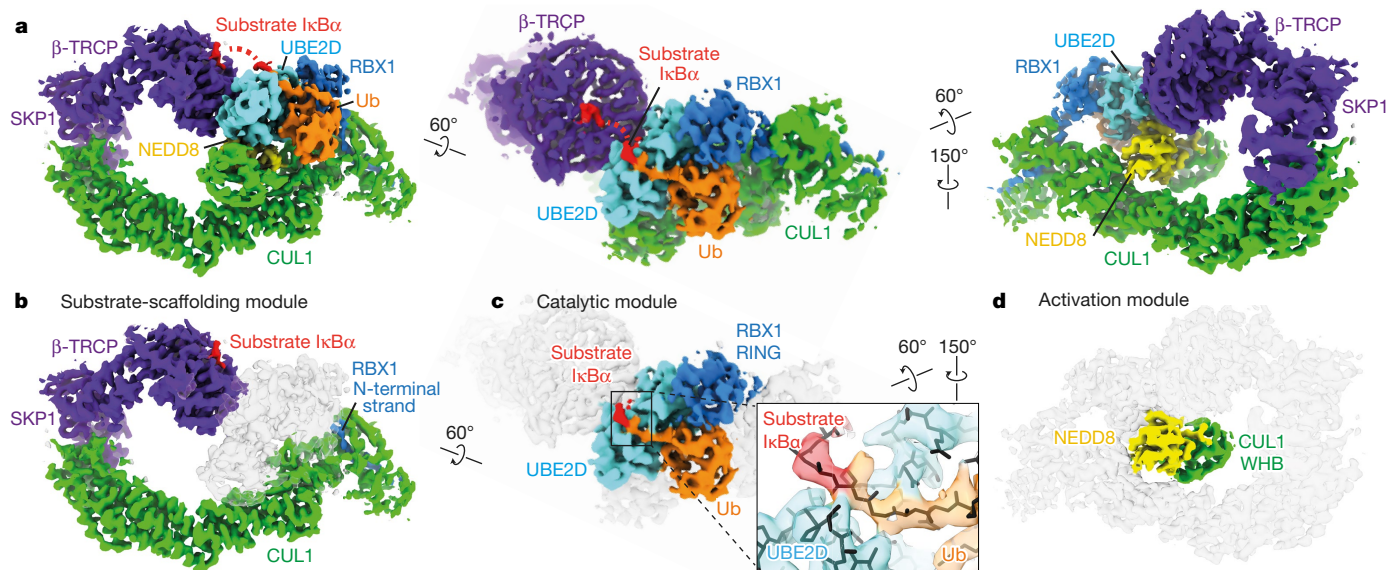


Fig. 2 | Cryo-EM structure representing neddylated CRL1^{β-TRCP}-mediated ubiquitin transfer from UBE2D to the substrate IkBα. **a**, Cryo-EM density representing the neddylated CRL1^{β-TRCP}-UBE2D-Ub-IkBα substrate intermediate, in which UBE2D-Ub is activated and juxtaposed with the substrate. **b**, The substrate-scaffolding module connects β-TRCP-bound

substrate to the intermolecular cullin-RBX (C/R) domain. **c**, The catalytic module consists of RING-UBE2D-Ub of RBX1 in the canonical closed activated conformation, and additional density corresponding to the chemical surrogate for the substrate undergoing ubiquitylation. **d**, NEDD8 and the covalently linked WHB domain of CUL1 form the activation module.

which is opposite the active site for ubiquitylation (Fig. 3c, Extended Data Fig. 7a). Contacts resemble those described for free NEDD8 or ubiquitin binding to the backside of UBE2D and allosterically stimulating the intrinsic reactivity of an isolated RING-UBE2D-Ub subcomplex^{35–38}. We examined intrinsic reactivity by monitoring ubiquitin discharge to free lysine using a previously described hyperactive neddylated CRL1^{β-TRCP} mutant³⁹ and high concentrations of enzyme and lysine. Substrate-independent ubiquitin transferase activity was impaired by a NEDD8 mutant that disrupts the integrity of the activation module, and by mutations of UBE2D that hinder its interactions with the covalently linked ubiquitin, the RING domain of RBX1 or NEDD8 (Extended Data Fig. 7c–f). The architecture observed in the neddylated CRL1^{β-TRCP}-UBE2D-Ub-substrate complex structure may therefore both stimulate the intrinsic reactivity of the UBE2D-Ub intermediate and place the catalytic centre in proximity to the β-TRCP-bound substrate.

The activation module is itself positioned by the binding of NEDD8 to the substrate-scaffolding module. Leu2, Lys4, Glu14, Asp16, Arg25, Arg29, Glu32, Gly63 and Gly64 of NEDD8—which nestle in a concave CUL1 surface—differ in ubiquitin, and these amino acids account for nearly one-third of the differences in sequence between the two proteins (Fig. 3d). The counterparts of these amino acids in ubiquitin would be expected to repel CUL1, which rationalizes the need for NEDD8 as a distinctive ubiquitin-like protein.

In addition, the catalytic module contacts both CUL1-RBX1 and the substrate receptor sides of the substrate-scaffolding module (Fig. 3e, Extended Data Fig. 5). On one side, the RING domain of RBX1 stacks on the Trp35 side chain of its C/R domain, which is consistent with previously reported effects of introducing a W35A mutation into RBX1³⁹. On the other side, the curved β-sheet of UBE2D complements the propeller of β-TRCP (Fig. 3e).

NEDD8 conformation coincidence coupling

Different conformations of ubiquitin and ubiquitin-like proteins have long been known to influence their interactions, with ubiquitin-binding domains selecting between ‘loop-in’ or ‘loop-out’ orientations of the

Leu8-containing β1/β2-loop. However, it remains largely unknown how these conformations might simultaneously affect binding to multiple partners⁴⁰. Our data show that NEDD8 must adopt the loop-out conformation to both form the activation module and to engage UBE2D in the catalytic module (Fig. 3f, Extended Data Fig. 7g–i). The conformation of NEDD8 apparently serves as a coincidence detector, coupling noncovalent binding to the linked WHB domain of CUL1 and to the catalytic module.

Synergistic catalytic assembly

We next assessed the importance of the structurally observed interfaces by testing the effects of mutations in relevant regions of the proteins. We used an attenuated pulse-chase assay format that qualitatively, but exclusively, monitors NEDD8-activated substrate priming by CRL1^{β-TRCP} and UBE2D (Extended Data Fig. 8). Mutations that were designed to destroy the activation module (NEDD8(Q40E)) or to hinder interactions between the activation and catalytic modules (NEDD8(I44A) or UBE2D(S22R)) substantially impaired substrate priming, as did swapping key NEDD8 residues at the interface with the substrate-scaffolding module with those found in ubiquitin. Moreover, although the structural basis for ubiquitylation by other CRLs requires further investigation, these mutations also impair the UBE2D-mediated priming of a cyclin E phosphopeptide substrate with neddylated CRL1^{FBW7}, and of IKZF zinc finger 2 by neddylated CRL4^{CRBN/Pomalidomide} (Extended Data Fig. 8g–o).

Given that the neddylated CRL1^{β-TRCP}-UBE2D-Ub-substrate intermediate depends on several conformational changes and large interfaces within and between modules (Fig. 2, Extended Data Figs. 5, 7), we hypothesized that pairing mutations that affect various interfaces would have synergistic effects. We used two types of experiment to define kinetic parameters for peptide substrate ubiquitylation—the Michaelis constant, K_m , was measured by titrating UBE2D, and the rate constant, k_{obs} , for ubiquitin transfer was determined by rapid quench-flow at saturating UBE2D concentrations to isolate catalytic defects. We drew three main conclusions from the results (Fig. 4a, Extended Data Fig. 1, Extended Data Table 1). First, although no

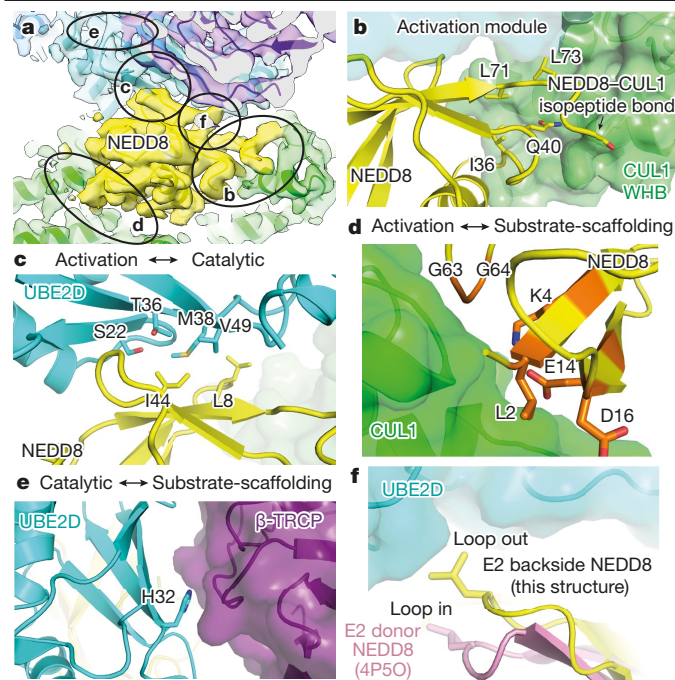


Fig. 3 | Intra- and inter-module interfaces specifying the catalytic architecture for ubiquitin priming of the substrate by neddylated CRL1^{β-TRCP} with UBE2D. **a**, Cryo-EM density highlighting noncovalent interfaces that contribute to the catalytic architecture for neddylated CRL1^{β-TRCP}-mediated ubiquitin transfer from UBE2D to a substrate. Circled regions correspond to interfaces within the activation module, and between activation and catalytic, activation and substrate-scaffolding, and catalytic and substrate-scaffolding modules shown in **b–f**. **b**, Close-up view of the intra-activation module interface, showing the buried polar residue Gln40 of NEDD8 and the Ile36/Leu71/Leu73 hydrophobic patch making noncovalent interactions with the WHB domain of CUL1, adjacent to the isopeptide bond linking NEDD8 and CUL1. **c**, Close-up view of the interface between the activation and catalytic modules, showing key residues at the interface between NEDD8 and the backside of UBE2D. **d**, Close-up view highlighting (in orange) the residues of NEDD8 that differ in ubiquitin, and that are at the interface with the substrate-scaffolding module. **e**, Close-up view highlighting His32 of UBE2D at the interface with the substrate-scaffolding module. **f**, Close-up view indicating the role of the loop-out conformation of NEDD8, which is required for the binding of UBE2D.

individual mutation is as detrimental as eliminating neddylation, the mutation of each interface on its own shows a decrease in k_{obs}/K_m of approximately tenfold or more compared with the wild-type. The most detrimental ‘single’ mutation is the substitution of NEDD8 with Ub(R72A)—this confirms the importance of NEDD8 and of interactions between the activation- and substrate-scaffolding modules. The next most detrimental mutation is Q40E in NEDD8—this underscores the importance of the structure of the activation module and its role in establishing the loop-out conformation of NEDD8. Second, combining mutations at several sites has a devastating effect on activity, even if the effect of the individual mutations alone is mild. For example, when using NEDD8(I44A) the activity is reduced by tenfold compared with the wild-type; however, when this mutant is combined with UBE2D(H32A) at the catalytic module–substrate receptor interface—which is mildly defective in an attenuated assay with subsaturating UBE2D (Extended Data Fig. 8f)—a near 200-fold reduction in activity is observed. Third, the mutations have comparatively little effect on chain elongation, which is consistent with the structure of the neddylated CRL1^{β-TRCP}–UBE2D–Ub–substrate intermediate defining ubiquitin transfer from UBE2D directly to the unmodified substrate.

Discussion

Our cryo-EM structure representing the neddylated CRL1^{β-TRCP}–UBE2D–Ub–substrate intermediate suggests a model for substrate priming that addresses many longstanding questions. First, rapid substrate ubiquitylation can be explained by NEDD8, the cullin and RBX1-bound UBE2D–Ub making numerous interactions that activate UBE2D and synergistically place the catalytic centre adjacent to β-TRCP. Second, biochemical features of neddylated CRLs that were incompatible with previous structures are now rationalized, including NEDD8-stimulated crosslinking between a CRL1^{β-TRCP}-bound phosphopeptide and UBE2D⁴; simultaneous NEDD8 linkage to a cullin and binding to the backside of RBX1-bound UBE2D^{3,7,37}; and the detrimental effects of bacterial-effector-catalysed NEDD8 Gln40 deamidation^{24–26} (Figs. 2, 3). Third, residues in ubiquitin that differ from those in NEDD8 would clash in the catalytic architecture (Fig. 3d, Extended Data Table 1), thus rationalizing the existence of NEDD8 as a distinct ubiquitin-like protein.

In the absence of other factors, the scaffolding module of neddylated CRL1^{β-TRCP} robustly bridges the substrate with the C/R domain, whereas NEDD8, its linked CUL1 WHB domain and the RING domain of RBX1 are relatively dynamic and apparently mobile (Extended Data Fig. 2). These mobile entities are harnessed in the neddylated CRL1^{β-TRCP}–UBE2D–Ub–substrate intermediate (Fig. 2). The numerous requisite protein–protein interactions and conformational changes suggest that there could be several routes to the catalytic architecture (Fig. 4b), in which the formation of interfaces successively narrows the range of options—akin to progression down a free-energy funnel. Because ubiquitylation does occur with mutant substrates or enzymes, albeit at substantially lower rates (Extended Data Table 1), we cannot exclude that ubiquitin could be transferred from RING- and NEDD8-bound UBE2D in various orientations relative to the substrate-scaffolding module. However, if the thioester bond is both in the RING-activated configuration and adjacent to the substrate—as in the structure—this would increase the rate at which the presumably random exploration of three-dimensional space by a substrate lysine would lead to productive collision with the active site. Accordingly, reducing any single contribution to the structurally observed catalytic architecture increases the relative importance of other contacts—even lesser ones (Fig. 4a, Extended Data Table 1).

Whereas neddylated CRL1^{β-TRCP} and UBE2D seem to be optimal for ubiquitin priming of peptide-like substrates, the limited effect of mutations on the linkage of subsequent ubiquitins (Extended Data Table 1) raises the possibility that different forms of ubiquitylation involve alternative—currently unknown—catalytic architectures. Although not overtly observed for our substrates with a single acceptor lysine, an outstanding question is whether there are other circumstances in which a substrate-linked ubiquitin could mimic NEDD8 and activate further ubiquitylation. Moreover, in addition to UBE2D, neddylated CRLs recruit a range of other ubiquitin-carrying enzymes—from ARIH-family RBR E3s for substrate priming to other E2s for polyubiquitylation^{21,41–43}—and UBXD7, which in turn recruits the AAA-ATPase p97 to process some ubiquitylated substrates⁴⁴. We speculate that these CRL partners uniquely harness the dynamic NEDD8, its linked CUL WHB domain and/or the RING domain of RBX1 to specify distinct catalytic activities, much like the thioester-linked UBE2D–Ub intermediate captures neddylated CRL1^{β-TRCP} through multiple surfaces to specify substrate priming. The malleability of neddylated CRLs—coupled with numerous ubiquitin carrying enzyme partners—may underlie successful molecular glue or PROTAC-targeted protein degradation, whereas potential limitations in their ability to achieve the optimal multivalent catalytic architectures may explain failures in such chemically directed ubiquitylation of heterologous substrates⁴⁵. Additionally, it seems likely that the conformational dynamics of unneddylated CRLs (Extended Data Fig. 2) would enable transitioning between the different conformations coordinating cycles of neddylation–deneddylation

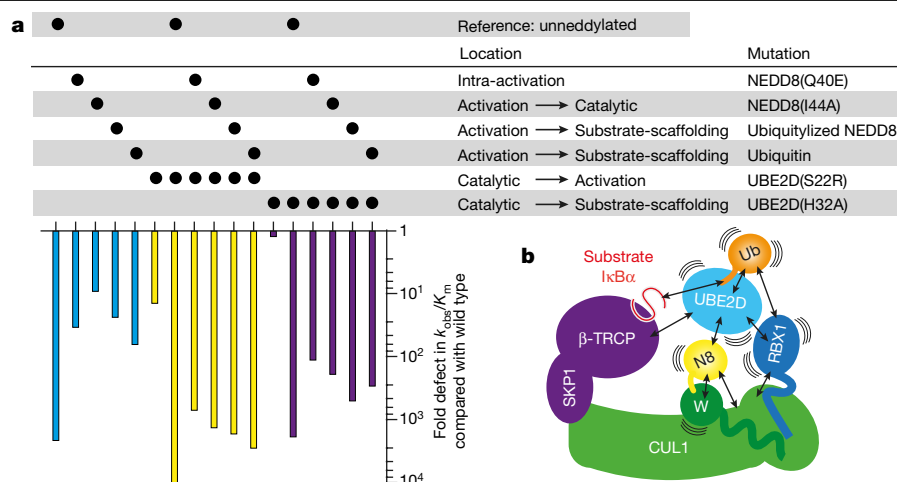


Fig. 4 | Multifarious interactions configuring rapid substrate priming. **a**, Effects of the indicated mutants—within the activation module, between activation and catalytic, activation and substrate-scaffolding and substrate-scaffolding and catalytic modules, alone or in combination—on the catalytic efficiency of substrate priming, as quantified by overall fold difference in k_{obs}/K_m compared with wild-type neddylation of CUL1^{β-TRCP} and UBE2D-catalysed ubiquitylation of a peptide substrate. Reactions with unneddylated CUL1^{β-TRCP} serve as a reference, and used CUL1(K720R) to prevent obscuring the interpretation of results by artefactual ubiquitin transfer to CUL1 and the resultant artefactual activation of substrate priming. Graphs show the average value from two different experiments (technical replicates), for which curve fits and values are provided in Extended Data Fig. 1 and Extended Data Table 1. **b**, On their own, neddylation of CUL1^{β-TRCP} and UBE2D-Ub are dynamic, and at an

extreme their constituent proteins and/or domains may be substantially mobile. Mobile entities are harnessed in the neddylation of CUL1^{β-TRCP}–UBE2D–Ub–substrate intermediate. There could be multiple routes to the catalytic architecture. It seems equally plausible that the UBE2D–Ub intermediate would first encounter either the RING domain of RBX1, or NEDD8—either of which would increase the effective concentration for the other interaction. Likewise, noncovalent-binding between NEDD8 and its linked WHB domain, or with the backside of UBE2D, would stabilize the loop-out conformation of NEDD8—this would also favour the other interaction. Ultimately, NEDD8, the cullin, and the RBX1-bound UBE2D–Ub intermediate make numerous interactions that synergistically establish a distinctive catalytic architecture that places UBE2D adjacent to β-TRCP.

with CAND1-driven substrate–receptor exchange^{39,46–49}. Thus, the multifarious nature of interactions and conformations that determine robust and rapid substrate priming—as revealed by the structure of the neddylation of CUL1^{β-TRCP}–UBE2D–Ub–substrate intermediate—also provides a mechanism by which common elements can be transformed by different protein partners to interconvert between distinct CRL assemblies to meet the cellular demand for ubiquitylation.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2000-y>.

- Lydeard, J. R., Schulman, B. A. & Harper, J. W. Building and remodelling Cullin–RING E3 ubiquitin ligases. *EMBO Rep.* **14**, 1050–1061 (2013).
- Read, M. A. et al. NedD8 modification of Cul-1 activates SCF^{β-TRCP}-dependent ubiquitination of IκBα. *Mol. Cell. Biol.* **20**, 2326–2333 (2000).
- Duda, D. M. et al. Structural insights into NEDD8 activation of cullin–RING ligases: conformational control of conjugation. *Cell* **134**, 995–1006 (2008).
- Saha, A. & Deshaies, R. J. Multimodal activation of the ubiquitin ligase SCF by NedD8 conjugation. *Mol. Cell* **32**, 21–31 (2008).
- Yamoa, K. et al. Autoinhibitory regulation of SCF-mediated ubiquitination by human cullin 1's C-terminal tail. *Proc. Natl Acad. Sci. USA* **105**, 12230–12235 (2008).
- Soucy, T. A. et al. An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature* **458**, 732–736 (2009).
- Zheng, N. et al. Structure of the Cul1–Rbx1–Skp1–F box^{Skp2} SCF ubiquitin ligase complex. *Nature* **416**, 703–709 (2002).
- Jin, J. et al. Systematic analysis and nomenclature of mammalian F-box proteins. *Genes Dev.* **18**, 2573–2580 (2004).
- Willems, A. R., Schwab, M. & Tyers, M. A hitchhiker's guide to the cullin ubiquitin ligases: SCF and its kin. *Biochim. Biophys. Acta* **1695**, 133–170 (2004).
- Angers, S. et al. Molecular architecture and assembly of the DDB1–CUL4A ubiquitin ligase machinery. *Nature* **443**, 590–593 (2006).
- Jin, J., Arias, E. E., Chen, J., Harper, J. W. & Walter, J. C. A family of diverse Cul4–Ddb1-interacting proteins includes Cdt2, which is required for S phase destruction of the replication factor Cdt1. *Mol. Cell* **23**, 709–721 (2006).

- Yu, C. et al. Gln40 deamidation blocks structural reconfiguration and activation of SCF ubiquitin ligase complex by NedD8. *Nat. Commun.* **6**, 10053 (2015).
- Pierce, N. W. et al. Cand1 promotes assembly of new SCF complexes through dynamic exchange of F box proteins. *Cell* **153**, 206–215 (2013).
- Stanley, D. J. et al. Inhibition of a NEDD8 cascade restores restriction of HIV by APOBEC3G. *PLoS Pathog.* **8**, e1003085 (2012).
- Yaron, A. et al. Identification of the receptor component of the IκBα-ubiquitin ligase. *Nature* **396**, 590–594 (1998).
- Winston, J. T. et al. The SCF^{β-TRCP}-ubiquitin ligase complex associates specifically with phosphorylated destruction motifs in IκBα and β-catenin and stimulates IκBα ubiquitination in vitro. *Genes Dev.* **13**, 270–283 (1999).
- Spencer, E., Jiang, J. & Chen, Z. J. Signal-induced ubiquitination of IκBα by the F-box protein Slimb/β-TRCP. *Genes Dev.* **13**, 284–294 (1999).
- Hart, M. et al. The F-box protein β-TRCP associates with phosphorylated β-catenin and regulates its activity in the cell. *Curr. Biol.* **9**, 207–211 (1999).
- Latres, E., Chiaur, D. S. & Pagano, M. The human F box protein β-TRCP associates with the Cul1/Skp1 complex and regulates the stability of β-catenin. *Oncogene* **18**, 849–854 (1999).
- Wu, G. et al. Structure of a β-TRCP1–Skp1–β-catenin complex: destruction motif binding and lysine specificity of the SCF^{β-TRCP} ubiquitin ligase. *Mol. Cell* **11**, 1445–1456 (2003).
- Wu, K., Kovacev, J. & Pan, Z. Q. Priming and extending: a UbH5/Cdc34 E2 handoff mechanism for polyubiquitination on a SCF substrate. *Mol. Cell* **37**, 784–796 (2010).
- Frescas, D. & Pagano, M. Deregulated proteolysis by the F-box proteins SKP2 and β-TRCP: tipping the scales of cancer. *Nat. Rev. Cancer* **8**, 438–449 (2008).
- Margottin, F. et al. A novel human WD protein, h-βTrCP, that interacts with HIV-1 Vpu connects CD4 to the ER degradation pathway through an F-box motif. *Mol. Cell* **1**, 565–574 (1998).
- Cui, J. et al. Glutamine deamidation and dysfunction of ubiquitin/NEDD8 induced by a bacterial effector family. *Science* **329**, 1215–1218 (2010).
- Jubelin, G. et al. Pathogenic bacteria target NEDD8-conjugated cullins to hijack host-cell signaling pathways. *PLoS Pathog.* **6**, e1001128 (2010).
- Morikawa, H. et al. The bacterial effector Cif interferes with SCF ubiquitin ligase function by inhibiting deneddylation of Cullin1. *Biochem. Biophys. Res. Commun.* **401**, 268–274 (2010).
- Tang, X. et al. Suprafacial orientation of the SCF^{Cdc4} dimer accommodates multiple geometries for substrate ubiquitination. *Cell* **129**, 1165–1176 (2007).
- Dou, H., Buetow, L., Sibbet, G. J., Cameron, K. & Huang, D. T. BIRC7–E2 ubiquitin conjugate structure reveals the mechanism of ubiquitin transfer by a RING dimer. *Nat. Struct. Mol. Biol.* **19**, 876–883 (2012).
- Plechanová, A., Jaffray, E. G., Tatham, M. H., Naismith, J. H. & Hay, R. T. Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. *Nature* **489**, 115–120 (2012).
- Pruneda, J. N. et al. Structure of an E3:E2–Ub complex reveals an allosteric mechanism shared among RING/U-box ligases. *Mol. Cell* **47**, 933–942 (2012).
- Brzovic, P. S. & Kleit, R. E. Ubiquitin transfer from the E2 perspective: why is UbH5 so promiscuous? *Cell Cycle* **5**, 2867–2873 (2006).
- Low, T. Y. et al. A systems-wide screen identifies substrates of the SCF^{β-TRCP} ubiquitin ligase. *Sci. Signal.* **7**, rs8 (2014).

33. Lumb, K. J. & Kim, P. S. A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled coil. *Biochemistry* **34**, 8642–8648 (1995).
34. Kawakami, T. et al. NEDD8 recruits E2-ubiquitin to SCF E3 ligase. *EMBO J.* **20**, 4003–4012 (2001).
35. Ozkan, E., Yu, H. & Deisenhofer, J. Mechanistic insight into the allosteric activation of a ubiquitin-conjugating enzyme by RING-type ubiquitin ligases. *Proc. Natl Acad. Sci. USA* **102**, 18890–18895 (2005).
36. Brzovic, P. S., Lissounov, A., Christensen, D. E., Hoyt, D. W. & Klevit, R. E. A UbcH5/ubiquitin noncovalent complex is required for processive BRCA1-directed ubiquitination. *Mol. Cell* **21**, 873–880 (2006).
37. Sakata, E. et al. Direct interactions between NEDD8 and ubiquitin E2 conjugating enzymes upregulate cullin-based E3 ligase activity. *Nat. Struct. Mol. Biol.* **14**, 167–168 (2007).
38. Buetow, L. et al. Activation of a primed RING E3-E2-ubiquitin complex by non-covalent ubiquitin. *Mol. Cell* **58**, 297–310 (2015).
39. Scott, D. C. et al. Structure of a RING E3 trapped in action reveals ligation mechanism for the ubiquitin-like protein NEDD8. *Cell* **157**, 1671–1684 (2014).
40. Hospenthal, M. K., Freund, S. M. & Komander, D. Assembly, analysis and architecture of atypical ubiquitin chains. *Nat. Struct. Mol. Biol.* **20**, 555–565 (2013).
41. Scott, D. C. et al. Two distinct types of E3 ligases work in unison to regulate substrate ubiquitylation. *Cell* **166**, 1198–1214.e24 (2016).
42. Huttenhain, R. et al. ARIH2 is a Vif-dependent regulator of CUL5-mediated APOBEC3G degradation in HIV infection. *Cell Host Microbe* **26**, 86–99.e7 (2019).
43. Hill, S. et al. Robust cullin–RING ligase function is established by a multiplicity of poly-ubiquitylation pathways. *eLife* **8**, e51163 (2019).
44. den Besten, W., Verma, R., Kleiger, G., Oania, R. S. & Deshaies, R. J. NEDD8 links cullin–RING ubiquitin ligase function to the p97 pathway. *Nat. Struct. Mol. Biol.* **19**, 511–516 (2012).
45. Schapira, M., Calabrese, M. F., Bullock, A. N. & Crews, C. M. Targeted protein degradation: expanding the toolbox. *Nat. Rev. Drug Discov.* **18**, 949–963 (2019).
46. Goldenberg, S. J. et al. Structure of the Cand1–Cul1–Roc1 complex reveals regulatory mechanisms for the assembly of the multisubunit cullin-dependent ubiquitin ligases. *Cell* **119**, 517–528 (2004).
47. Mosadeghi, R. et al. Structural and kinetic analysis of the COP9-signalosome activation and the cullin–RING ubiquitin ligase deneddylation cycle. *eLife* **5**, e12102 (2016).
48. Cavadini, S. et al. Cullin–RING ubiquitin E3 ligase regulation by the COP9 signalosome. *Nature* **531**, 598–603 (2016).
49. Liu, X. et al. Cand1-mediated adaptive exchange mechanism enables variation in F-box protein expression. *Mol. Cell* **69**, 773–786.e6 (2018).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Methods

Cloning, protein expression and purification

All proteins are of human origin. All variants of UBE2D2, UBE2D3, UBE2M, RBX1, CUL1, CUL4, NEDD8 and ubiquitin were generated using PCR, Quikchange (Agilent), or were synthesized by Twist Biosciences.

UBE2D2 was purified as previously described⁵⁰, and UBE2D3 was purified in a similar manner. Ubiquitin was expressed in BL21(DE3) RIL as previously described⁵¹. Wild-type CUL1, RBX1(5-C), SKP1, β -TRCP2, CUL4A (from residue 38 to C terminus, hereafter referred to as CUL4), CRBN, DDB1 and UBA1 were cloned into pLIB vectors⁵². GST-TEV-RBX1 and CUL1, GST-TEV-RBX1 and CUL4A, His-TEV- β -TRCP2 and SKP1 or His-TEV-DDB1 and GST-TEV-CRBN were co-expressed by co-infecting with two baculoviruses. UBA1 was cloned with an N-terminal GST-tag with a TEV cleavage site. These proteins were expressed in *Trichoplusia ni* High-Five insect cells, purified by either GST or nickel-affinity chromatography, overnight TEV cleavage, followed by ion-exchange and size-exclusion chromatography. All variants of CUL1-RBX1 were purified similarly. Purification of NEDD8, UBE2M, APPBP1-UBA3, SKP1-FBW7 (from residue 263 to C terminus⁵³), neddylation of CUL1-RBX1, and fluorescent labelling of ubiquitin used for biochemical assays were performed as previously described³⁹. β -TRCP1 (monomeric form, from residue 175 to C terminus²⁰, hereafter referred to as β -TRCP1) with an N-terminal His-MBP followed by a TEV cleavage site was cloned into a pRSFDuet vector with SKP1 $\Delta\Delta$ (SKP1 with two internal deletions, of residues 38–43 and 71–82)⁵⁴. SKP1 $\Delta\Delta$ - β -TRCP1 was expressed in BL21(DE3) Gold *E. coli* at 18 °C, purified with nickel affinity chromatography, followed by TEV cleavage, anion exchange and size-exclusion chromatography. Modification of RBX1-CUL1 and RBX1-CUL4A by ubiquitin instead of NEDD8 was performed with the Ub(R72A) mutant that allows its activation and conjugation by neddylation enzymes APPBP1-UBA3 and UBE2M^{55,56}. The reaction for ubiquitylating CUL4A-RBX1 was performed at pH 8.8 to drive the reaction to completion. The previously described Y130L mutant of UBE2M was used to modify CUL1-RBX1 with the I44A mutant of NEDD8³⁹. IKZF1 ZF2 (residues 141–169, with two point mutations (K157R/K165R) and with a lysine added at position 140 to create a single target lysine at the N terminus⁵⁷) was cloned with an N-terminal GST with a 3C-Prescission cleavage site and a noncleavable C-terminal Strep-tag. IKZF1 ZF2 was purified by GST affinity chromatography, 3C-Prescission cleavage overnight, and size-exclusion chromatography. UBE4B RING-like U-box domain (residues 1200–C terminus) containing D1268T and N1271T point mutations that enhance activity⁵⁸ (hereafter referred to as UBE4B) was cloned with an N-terminal GST with TEV cleavage site. UBE4B was purified by GST affinity chromatography, TEV cleavage overnight, followed by ion exchange and size-exclusion chromatography.

Peptides

All peptides were stated to be of >95% purity by HPLC and were used as received.

Peptides used to quantify enzyme kinetics had the following sequences: IkB α , KERLLDDRHD(pS)GLD(pS)MRDEERRASY (obtained from New England Peptide); β -catenin short, KSYLD(pS)GIH(pS)GATTAPRRASY (obtained from Max Planck Institute of Biochemistry Core Facility); β -catenin medium, KAWQQSYLD(pS)GIH(pS)GATTAPRRASY (obtained from New England Peptide); β -catenin long, KAAVSHWQQSYLD(pS)GIH(pS)GATTAPRRASY (obtained from Max Planck Institute of Biochemistry Core Facility); β -catenin for sortase-mediated transpeptidation to ubiquitin to generate a homogeneously ubiquitin-linked substrate, GGGGYLD(pS)GIH(pS)GATTAPRRASY (obtained from Max Planck Institute of Biochemistry Core Facility).

Peptides used for the qualitative assays monitoring substrate priming—that is, fluorescent ubiquitin transfer from UBE2D-Ub to substrate—had the following sequences: IkB α , KERLLDDRHD(pS)GLD(pS)

MKDEE (as previously described⁴¹); CyE, KAMLSEQNRSPLPSGLL(pT)PPQ(pS)GRRASY (as previously described⁴¹).

The nonmodifiable substrate analogue used in the competition experiment in Extended Data Fig. 3d had the following sequence: IkB α , RRERLLDDRHD(pS)GLD(pS)MRDEE (obtained from Max Planck Institute of Biochemistry Core Facility).

Peptides used in cryo-EM experiments are as follows: for the structure representing neddylation CRL1 ^{β -TRCP}-UBE2D-Ub-IkB α substrate described in detail and cryo-EM experiments shown in Extended Data Fig. 3e, f, h, i: IkB α , CKKERLLDDRHD(pS)GLD(pS)MKDEEDYKDDDDK (obtained from Max Planck Institute of Biochemistry Core Facility); for cryo-EM reconstructions of unneddylation and neddylation CRL1 ^{β -TRCP}-IkB α substrate shown in Extended Data Fig. 2a, b: IkB α , KKERLLDDRHD(pS)GLD(pS)MKDEE (as previously described⁴¹).

Enzyme kinetics

UBE2D3 titrations under substrate single-encounter conditions for estimation of the K_m for E2 used by neddylation CRL1. These experiments used full-length CRL1 ^{β -TRCP2} and UBE2D3, referred to here as CRL1 ^{β -TRCP} and UBE2D. Fifty micromolar peptide substrate (for a list of peptides that were used in the assay, see ‘Peptides’) was radiolabelled with 5 kU of cAMP-dependent protein kinase (New England Biolabs) in the presence of [γ -³²P]ATP for 1 h at 30 °C. Two mixtures were prepared before initiation of the reaction: a UBA1/Ub mix containing unlabelled substrate competitor peptide that was identical in sequence to the labelled one (the one exception being the Ub- β -catenin substrate, in which the unlabelled β -catenin for sortase peptide was used); and a neddylation or unneddylation (with the CUL1(K720R) mutant, to prevent low-level ubiquitylation by UBE2D³) CRL1 ^{β -TRCP}/labelled peptide substrate mix. The UBA1/Ub mix contained reaction buffer composed of 30 mM Tris-HCl, 100 mM NaCl, 5 mM MgCl₂, 2 mM ATP and 2 mM DTT pH 7.5. The concentration of ubiquitin was 80 μ M, with 1 μ M UBA1 and 100 μ M unlabelled peptide. UBE2D was first prepared as a twofold dilution series from a variable stock concentration, then introduced individually into tubes containing equal amounts of the UBA1/Ub mix. The CRL1 ^{β -TRCP}/labelled peptide substrate mix contained the same reaction buffer as the UBA1/Ub/UBE2D mix, 0.5 μ M CRL1 ^{β -TRCP}, and 0.2 μ M labelled peptide substrate. The reactions were initiated at 22 °C by combining equal volumes of both mixes, rapidly vortexed and quenched after 10 s in 2 \times SDS-PAGE buffer containing 100 mM Tris-HCl, 20% glycerol, 30 mM EDTA, 4% SDS and 4% β -mercaptoethanol pH 6.8. Each titration series was performed in duplicate and resolved on hand-cast, reducing 18% SDS-PAGE gels. The gels were imaged on a Typhoon 9410 Imager and quantification of substrate and products was performed using Image Quant (GE Healthcare). The product of each lane was measured as the fraction of the ubiquitylated products divided by the total signal, plotted against the UBE2D concentration, and fit to the Michaelis–Menten equation to estimate K_m (GraphPad Prism software). The standard error was calculated using Prism and has been provided in Extended Data Table 1 for all estimates of K_m .

Estimating the rates of ubiquitin transfer to CRL1-bound substrate using pre-steady-state kinetics. These experiments used full-length CRL1 ^{β -TRCP2} and UBE2D3, referred to here as CRL1 ^{β -TRCP} and UBE2D, respectively. Separate UBA1/UBE2D/Ub and CRL1 ^{β -TRCP}/labelled peptide substrate mixes were prepared to assemble single-encounter ubiquitylation reactions. For most reactions, the UBA1/UBE2D3/Ub mix contained reaction buffer, 80 μ M Ub, 1 μ M UBA1, 40 μ M UBE2D, and 200 μ M unlabelled competitor substrate peptide. For all reactions containing either CUL1(K720R) or UBE2D(H32A) assayed with wild-type neddylation CUL1, CUL1 modified either with the mutant NEDD8(I44A) or with Ub(R72A) permitting ligation to CUL1, 120 μ M Ub and 70 μ M UBE2D were used. The CRL1 ^{β -TRCP}/labelled peptide substrate mixes contained reaction buffer, 0.5 μ M CRL1 ^{β -TRCP}, and 0.2 μ M labelled peptide (for a list of peptides that were used in the assay, see ‘Peptides’).

Article

Each mix was separately loaded into the left or right sample loops on a KinTek RQF-3 quench flow instrument, and successive time points were taken at 22 °C by combining the mixtures with drive buffer composed of 30 mM Tris-HCl and 100 mM NaCl pH 7.5. Reactions were quenched at various time points in 2× SDS–PAGE buffer to generate the time courses. Substrate and products from each time point were resolved on hand-cast, reducing 18% SDS–PAGE gels. The gels were imaged on a Typhoon 9410 Imager, and substrate and product bands were individually quantified as a percentage of the total signal for each time point using ImageQuant (GE Healthcare). Reactions were performed in duplicate, and the average of each substrate or product band was used for the analysis. The data for substrate (S0) or mono-ubiquitylated product (S1) bands were fit to their respective closed-form solutions as previously described⁵⁹ using Mathematica to obtain the values $k_{\text{obs}}^{\text{S0-S1}}$ and $k_{\text{obs}}^{\text{S1-S2}}$ (Extended Data Table 1). The standard errors were calculated in Mathematica and have been provided in Extended Data Table 1 for all estimates of k_{obs} .

Multiturnover assays with short β -catenin. The multiturnover assay showing the ubiquitin transfer to short β -catenin (Extended Data Fig. 6c, d) was performed as described in ‘Estimating the rates of ubiquitin transfer to CRL1-bound substrate using pre-steady-state kinetics’, but without the excess unlabelled β -catenin peptide substrate. Time points were collected by quenching in 2× SDS–PAGE loading buffer. Substrate and product were separated by SDS–PAGE followed by autoradiography. The fraction of unmodified substrate (S0) was quantified and fit to either a one-phase decay (unneddylated) or linear (neddylated) model (Prism 8). Similarly, products containing five or more ubiquitins were quantified and fit to either an exponential growth (neddylated) or linear (unneddylated) model. Experiments were performed in duplicate.

Generation of ubiquitylated β -catenin fusion via sortase reaction. A mimic of ubiquitylated β -catenin was generated by fusing a ubiquitin with a C-terminal LPETGG with a GGGG- β -catenin peptide. The reaction was incubated with concentrations of 50 μM UB^{LPETGG}, 300 μM GGGG- β -catenin peptide, and 10 μM 6×His-Sortase A for 10 min in 50 mM Tris, 150 mM NaCl, 10 mM CaCl₂, pH 8.0. Sortase A was removed by retention on nickel resin, and the product was further purified by size-exclusion chromatography in 25 mM HEPES, 150 mM NaCl, 1 mM DTT at pH 7.5.

The sequence of ubiquitin with sortase motif was as follows: MQIFVK-TLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLLIFAGKQLEDGRTLS-DYNIQKESTLHLVLRRLRGSGSLPETGG.

Other biochemical assays

For experiments comparing the activity of neddylated with unneddylated CRLs or CUL–RBX1 complexes, in the unneddylated versions the NEDD8 modification sites of CUL1 and CUL4 were mutated to Arg (CUL1(K720R) and CUL4A(K705R)) to prevent obscuring interpretation of the results by low-level ubiquitylation of the NEDD8 consensus Lys during the ubiquitylation reactions³.

Substrate priming assays. Experiments in Extended Data Fig. 8a–f used full-length CRL1 ^{β -TRCP2} and UBE2D3. Ubiquitylation of IkB α by CRL1 ^{β -TRCP} was monitored using a pulse-chase format that specifically detects CRL1 ^{β -TRCP}-dependent ubiquitin modification from UBE2D to IkB α independently of effects on UBA1-dependent formation of the UBE2D3–Ub intermediate. The pulse reaction generated a thioester-linked UBE2D–Ub intermediate and contained 10 μM UBE2D, 15 μM fluorescent ubiquitin, 0.2 μM UBA1 in 50 mM Tris, 50 mM NaCl, 2.5 mM MgCl₂, 1.5 mM ATP pH 7.5 incubated at room temperature for 10 min. The pulse reaction was quenched with 25 mM EDTA on ice for 5 min, then further diluted to 100 nM UBE2D in 25 mM MES, 150 mM NaCl pH 6.5 for subsequent mixture with components of the reaction for neddylated CRL1 ^{β -TRCP}-dependent ubiquitin transfer to the

substrate in the chase reaction. The chase reaction mix consisted of 400 nM CRL (NEDD8–CUL1–RBX1–SKP1– β -TRCP), and 1 μM substrate (phosphorylated peptide derived from IkB α) in 25 mM MES, 150 mM NaCl pH 6.5 incubated on ice. After the quench, the pulse reaction mix was combined with the chase reaction mix at a 1:1 ratio on ice. The final reaction concentrations were 50 nM UBE2D (in thioester-linked UBE2D–Ub complex) and 200 nM neddylated CRL1 ^{β -TRCP} to catalyse substrate ubiquitylation. Samples were taken at each time point, quenched with 2× SDS–PAGE sample buffer, protein components were separated on nonreducing SDS–PAGE, and the gel was scanned on an Amersham Typhoon imager (GE Healthcare).

Substrate priming reactions assaying the effects of variations in UBE2D shown in Extended Data Fig. 8g–i on CRL1 ^{β -TRCP}-dependent ubiquitylation—a phosphopeptide derived from CyE—were performed similarly to those for CRL1 ^{β -TRCP} as described above with 100 nM UBE2D–Ub (based on concentration of UBE2D from the pulse reaction), 500 nM neddylated CUL1–RBX1–SKP1–FBW7 (residues 263 to the C terminus), and 2.5 μM CyE phosphopeptide in 25 mM HEPES, 150 mM NaCl pH 7.5 at room temperature. Experiments testing the effects of variations in NEDD8 (or its substitution with Ub(R72A)) were performed in the same manner, except with 250 nM NEDD8 (or variant)-modified CUL1–RBX1–SKP1–FBW7 (residues 263 to the C terminus).

Our assay for CRL4 ^{CRBN} ubiquitylation of IKZF was established on the basis of findings that ZF2 mediates tight immunomodulatory-drug-dependent interactions sufficient to target degradation, and UBE2D3 contributes to the stability of CRL4 ^{CRBN} neomorphic substrates in cells^{57,60,61}. Substrate priming reactions showing controls and assaying effects of variations in NEDD8 and UBE2D3 shown in Extended Data Fig. 8j–o monitored ubiquitylation of IKZF ZF2 with 400 nM UBE2D–Ub (concentration determined by that of UBE2D in the chase reaction), 500 nM NEDD8–CUL4–RBX1–DDB1–CRBN, 5 μM pomalidomide and 2.5 μM IKZF ZF2 in 25 mM HEPES, 150 mM NaCl pH 7.5 at room temperature. Effects of swapping NEDD8 for Ub(R72A) on the CRL4 ^{CRBN} -mediated ubiquitylation of IKZF ZF2 are shown in Extended Data Fig. 8m and were performed similarly but with 100 nM UBE2D–Ub, 250 nM NEDD8 or Ub-modified CUL4–RBX1–DDB1–CRBN, 2.5 μM pomalidomide and 1.25 μM IKZF ZF2.

Assays for intrinsic activation of UBE2D–Ub intermediate. Assays shown in Extended Data Figs. 3b, g, 7d–f—monitoring neddylated CUL1–RBX1 activation of the thioester-linked UBE2D–Ub intermediate (that is, in the absence of substrate)—used the RBX1(N98R) variant that is hyperactive towards UBE2D–Ub³⁹. Experiments in Extended Data Figs. 3b, g, 7f were performed in pulse-chase format similar to substrate-priming assays, but with 9 μM UBE2D–Ub (loading reaction with 20 μM UBE2D, 30 μM Ub and 0.5 μM UBA1), 500 nM E3 and 5 mM free lysine. For unneddylated CUL1–RBX1- or UBE4B-dependent discharge, 50 mM free lysine was used instead. Discharge assays shown in Extended Data Fig. 7d, e used 5 μM UBE2D–Ub (loading reaction with 20 μM UBE2D, 20 μM Ub and 0.5 μM UBA1), 500 nM E3 and 10 mM free lysine for neddylated CUL1–RBX1-dependent discharge, and 50 mM free lysine for unneddylated CUL1–RBX1-dependent discharge. All assays were visualized by Coomassie-stained SDS–PAGE.

Generation of a stable proxy for the UBE2D–Ub–substrate intermediate

Preparation of His-TEV-Ub(1–75)-MESNa. His-TEV-Ub(1–75) was cloned using a previously described method⁶² into pTXB1 (New England Biolabs) and transformed into BL21(DE3) RIL. Cells were grown in terrific broth at 37 °C to an optical density at 600 nm (OD₆₀₀) of 0.8 and then induced with IPTG (0.5 mM), shaking overnight at 16 °C. The collected cells were resuspended (20 mM HEPES, 50 mM NaOAc, 100 mM NaCl, 2.5 mM PMSF pH 6.8), sonicated and then centrifuged (50,000g, 4 °C, 30 min). Ni-NTA resin (1 millilitre resin per litre of broth, Sigma Aldrich) was equilibrated with the resuspension buffer and incubated

with the cleared lysate at 4 °C on a roller (30 rpm) for 1 h. The resin was then transferred to a gravity column and washed (5 × 1 column volume with 20 mM HEPES, 50 mM NaOAc, 100 mM NaCl pH 6.8). Protein was then eluted (5 × 1 column volume with 20 mM HEPES, 50 mM NaOAc, 100 mM NaCl 300 mM imidazole pH 6.8). Ubiquitin was then cleaved from the chitin-binding domain by diluting the eluted protein 10:1 (v/v) with 20 mM HEPES, 50 mM NaOAc, 100 mM NaCl, 100 mM sodium 2-mercaptoethanesulfonate (Sigma Aldrich) pH 6.8. This solution was incubated at room temperature overnight on a roller (30 rpm). Ub-MESNa was finally purified by size-exclusion chromatography (SD75 HiLoad, GE Healthcare) equilibrated with 12.5 mM HEPES, 25 mM NaCl pH 6.5.

Sequence of His-TEV-Ub(1-75)-chitin-binding domain:

MGSSHHHHHHENLYFGSGGMQIFVKLTGKTITLEVEPSDTIEN-
VKAKIQD
KEGIPPDQQLIFAGKQLEDGRTLSDYNIQKESTLHLVRLRGCF-
AKGTNVL
MADGSIECIENIEVGKVMGKDGRPREVIKLPGRRETMYSVVQSKQH-
RAH
KSDSSREVPPELLKFTCNATHELVVRTPRSVRRLSRTIKGVEYFEVIT-
FEMGQ
KKAPDG

Native chemical ligation to make Ub(1-75)-Cys-IkBα. His-Ub(1-75)-MESNa (200 μM final concentration) and freshly dissolved IkBα peptide (H-CKKERLLDDRHDpSGLDpSMKDEEDYKDDDDK-OH) (1,000 μM final concentration) were combined in a 1.5-ml tube in 50 mM NaPO₄, 50 mM NaCl pH 6.5. This was incubated with rocking at 30 rpm for 1 h at room temperature before TCEP was added to 1 mM. After rocking for an additional hour at room temperature, the reaction was quenched by adding 500 mM NaPO₄ pH 8.0 to 45 mM. The entire solution was then incubated with Ni-NTA resin (300 μl for a 1 ml reaction) at 30 rpm for 1 h at 4 °C. In a gravity column, the resin was then washed with 6 × 300 μl 50 mM NaPO₄, 50 mM NaCl, 1 mM β-mercaptoethanol pH 8.0. Protein was eluted with 50 mM NaPO₄, 50 mM NaCl, 1 mM β-mercaptoethanol, 300 mM imidazole pH 8.0. Fractions were analysed by SDS-PAGE and nanodrop.

Formation of disulfide linkage between UBE2D Cys85 and Ub(1-75)-Cys-IkBα. The same approach was used to generate complexes for UBE2D2 and UBE2D3, referred to collectively as UBE2D. UBE2D(C21I/C107A/C111D) was purified from size-exclusion chromatography (see 'Cloning, protein expression and purification') and then immediately used without freezing. After size-exclusion chromatography, the protein was concentrated (Amicon, EMD Millipore) to 600 μM. Protein (2 × 100 μl) was separately desalted (2 × Zeba, 0.5 ml column, 7,000 molecular weight cut-off filter, Thermo Fisher) to 20 mM HEPES, 250 mM NaCl, 5 mM EDTA pH 7.0. Elutions were combined and immediately added together to 34 μl 10 mM 5,5'-dithiobis-(2-nitrobenzoic acid) (Sigma Aldrich, dissolved in 50 mM NaPO₄ pH 7.5) and mixed by pipetting before incubating at room temperature for 30 min. The solution was then desalted (2 × Zeba, 0.5 ml column, 7,000 molecular weight cut-off filter, Thermo Fisher) to 20 mM HEPES, 250 mM NaCl, 5 mM EDTA pH 7.0 at the same time that Ub(1-75)-Cys-IkBα (500 μl at 100 μM) was desalted (1 × Zeba, 2 ml column, 7,000 molecular weight cut-off filter, Thermo Fisher) to the same buffer. The UBE2D and ubiquitin components were then immediately combined and incubated at room temperature for 30 min, at which point the sample was loaded to a Superdex 75 Increase column (GE Healthcare) equilibrated with 20 mM HEPES, 250 mM NaCl, 5 mM EDTA pH 7.0.

Comparing the ability of stable proxy for the UBE2D-Ub-substrate intermediate and subcomplexes to compete with ubiquitylation. Assays comparing ubiquitylation in the presence of competitors (stable proxy for the UBE2D-Ub-substrate intermediate, stable isopeptide-linked mimic of UBE2D-Ub, and nonmodifiable substrate peptide) were carried out similarly to that described for our substrate priming

assay described in 'Substrate priming assays' with the following modifications. The assay was performed in pulse-chase format to exclude the potential for competitors to affect generation of the UBE2D-Ub intermediate. In the pulse reaction, a thioester-linked UBE2D-Ub intermediate was generated by incubating 10 μM UBE2D, 15 μM fluorescent Ub, and 0.2 μM UBA1 in a buffer that contained 50 mM Tris, 50 mM NaCl, 2.5 mM MgCl₂, and 1.5 mM ATP pH 7.6 at room temperature for 10 min. The pulse reaction was next quenched by the addition of an equal volume of 50 mM Tris, 50 mM NaCl, 50 mM EDTA pH 7.6 and placed on ice for 5 min, then further diluted to 100 nM in a buffer containing 25 mM MES pH 6.5 and 150 mM NaCl. The E3-substrate mix consisted of 400 nM NEDD8-CUL1-RBX1, 400 nM SKP1-β-TRCP, 1 μM IkBα peptide, with or without 1 μM competitor in a buffer consisting of 25 mM MES and 150 mM NaCl pH 6.5, and was incubated at 4 °C for 10 min to achieve equilibrium. Reactions were initiated on ice by the addition of an equal volume of pulse reaction to the E3-substrate mix, resulting in final reaction conditions of 50 nM UBE2D-Ub, 200 nM E3 neddylation CRL1^{β-TRCP}, and 500 nM substrate with or without 500 nM competitor. Samples were taken at the indicated time points and quenched with 2 × SDS-PAGE sample buffer. Substrate and products were then separated by SDS-PAGE, and subsequently visualized using an Amersham Typhoon imager (GE Healthcare).

Early attempt to visualize ubiquitin transfer by neddylation CRL1^{β-TRCP} and UBE2D and rationale for approaches to improve electron microscopy samples. In our initial attempt to determine a structure visualizing ubiquitin transfer by neddylation CRL1^{β-TRCP} and UBE2D, we used a full-length β-TRCP (β-TRCP2), which is a homodimer²⁷, and a proxy for a UBE2D-Ub-substrate intermediate based on a method used to capture a Sumoylation intermediate⁶³. In the previous study, SUMO was installed via an isopeptide bond on a residue adjacent to the E2 catalytic Cys, and substrate was crosslinked to the E2 Cys via an ethanedithiol linker. Here, we introduced the corresponding lysine substitution in the background of an optimized E2 (UBE2D2(L119K/C21I/C107A/C111D)). Using high concentrations of UBA1 and high pH, we generated an isopeptide-bonded complex between ubiquitin and this UBE2D2 variant in a manner dependent on the L119K mutation, and crosslinked the Cys of the IkBα substrate mimic peptide to the UBE2D-Ub complex using EDT as previously described⁶³. The resultant cryo-EM data map, shown in Extended Data Fig. 3e, presented two major challenges. First, the dimer exacerbated structural heterogeneity, with minor differences presumably based on natural motions between the two protomers. Second, the donor ubiquitin was poorly visible, presumably owing to it not being linked to the catalytic Cys. Thus, we generated many samples in parallel to overcome these challenges by (1) using a monomeric version of β-TRCP1 that had previously been crystallized²⁰; (2) removing two loops in SKP1 that are known to be flexible and not required for ubiquitylation activity (although they are required for CAND1-mediated substrate-receptor exchange)^{13,54}; (3) devising a chemical approach to synthesize a proxy for the UBE2D-Ub-substrate intermediate in which all three entities are simultaneously linked to the E2 catalytic Cys; (4) using a point mutant version of RBX1 that is hyperactive for substrate priming with UBE2D but defective for ubiquitin chain elongation with UBE2R-family E2s³⁹. Notably, with a *K_m* for UBE2D3 of 350 nM and rates of ubiquitylating the medium β-catenin peptide substrate of 8.9 s⁻¹ (S0-S1) and 0.25 s⁻¹ (S1-S2), we confirmed that the monomeric version of neddylation CRL1^{β-TRCP1} is kinetically indistinguishable from full-length, homodimeric neddylation CRL1^{β-TRCP2}.

Cryo-EM

Sample preparation. For neddylation or unneddylation CUL1-RBX1-SKP1-β-TRCP-IkBα samples, subcomplexes were mixed in an equimolar ratio with 1.5-fold excess substrate peptide, incubated for 30 min on ice and purified by size-exclusion chromatography in 25 mM HEPES, 150 mM NaCl, 1 mM DTT pH 7.5. The complex was further concentrated

Article

and crosslinked by GraFix⁶⁴. The sample was next liberated of glycerol using Zeba Desalt Spin Columns (Thermo Fisher), concentrated to 0.3 mg ml⁻¹, and 3 µl of sample was applied to R1.2/1.3 holey carbon grids (Quantifoil) and was plunge-frozen by Vitrobot Mark IV in liquid ethane. Structure determination of the neddylated CUL1-RBX1-SKP1-β-TRCP-Ub-UBE2D-IkBα complex used a similar method as above, with 1.5-fold excess of the stable proxy for the UBE2D-Ub-IkBα intermediate, but with no DTT in buffer. After SEC, GraFix, and desalting, 3 µl of 0.08 mg ml⁻¹ sample was applied to graphene oxide-coated Quantifoil R2/1 holey carbon grids (Quantifoil)⁶⁵ and was plunge-frozen by Vitrobot Mark IV in liquid ethane.

Electron microscopy. Datasets were collected on a Glacios cryo transmission electron microscope at 200 kV using a K2 Summit direct detector in counting mode. For the CUL1-RBX1-SKP1-β-TRCP1ΔD-IkBα dataset, 6,433 images were recorded at 1.181 Å per pixel with a nominal magnification of 36,000×. A total dose of 60 e⁻ Å⁻² was fractionated over 50 frames, with a defocus range of -1.2 µm to -3.3 µm. For NEDD8-CUL1-RBX1-SKP1-β-TRCP1ΔD-IkBα, 2,061 images were recorded at 1.885 Å per pixel with a nominal magnification of 22,000×. A total dose of 59 e⁻ Å⁻² was fractionated over 38 frames, with a defocus range of -1.2 µm to -3.3 µm.

Datasets were also collected on a Talos Arctica at 200 kV using a Falcon II direct detector in linear mode. For each sample, around 800 images were recorded at 1.997 Å per pixel with a nominal magnification of 73,000×. A total dose of approximately 60 e⁻ Å⁻² was fractionated over 40 frames, with a defocus range of -1.5 µm to -3.5 µm.

High-resolution cryo-EM data were collected on a Titan Krios electron microscope at 300 kV with a Quantum-LS energy filter, using a K2 Summit direct detector in counting mode. 9,112 images were recorded at 1.06 Å per pixel with a nominal magnification of 130,000×. A total dose of 70.2 e⁻ Å⁻² was fractionated over 60 frames, with a defocus range of -1.2 µm to -3.6 µm.

Data processing. Frames were motion-corrected using RELION-3.0⁶⁶ with dose weighting. Contrast transfer function was estimated using CTFFIND⁶⁷. Particles were picked with Gautomatch (K. Zhang, MRC Laboratory of Molecular Biology). Two-dimensional classification was performed in RELION-3.0, followed by 3D ab initio model building by sxvipier.py from SPARX⁶⁸. The initial model from sxvipier.py was imported to RELION-3.0 for further 3D classification, refinement, post-processing and particle polishing using frames 2–25.

Protein identification and model building. The final reconstructions displayed clear main chain and side chain densities, which enabled us to model and refine the atomic coordinates. Known components (CUL1-RBX1, RCSB Protein Data Bank codes (PDB) 1LDJ and 4P5O; SKP1ΔΔ-β-TRCP1, PDB 1P22 and 6M90; UBE2D-Ub with a backside-bound ubiquitin to be replaced by NEDD8 sequences, PDB 4V3L) were manually placed as a whole or in parts and fit with rigid-body refinement using UCSF Chimera⁶⁹. The resultant complete structure underwent rigid-body refinements in which each protein or domain was allowed to move independently. Further iterative manual model building and real space refinements were carried out until good geometry and map-to-model correlation was reached. Manual model building and rebuilding were performed using COOT⁷⁰, and Phenix.refine⁷¹ was used for real space refinement.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The atomic coordinates and electron microscopy maps have been deposited in the PDB with accession code 6TTU and the Electron

Microscopy Data Bank with codes EMD-10585, EMD-10578, EMD-10579, EMD-10580, EMD-10581, EMD-10582 and EMD-10583. Uncropped gel source data are included as Supplementary Information. All other reagents and data (for example, raw gels of replicate experiments and raw movie electron microscopy data) are available from the corresponding author upon request.

50. Kamadurai, H. B. et al. Insights into ubiquitin transfer cascades from a structure of a Ub-H5B-ubiquitin-HECT^{NEDD8} complex. *Mol. Cell* **36**, 1095–1102 (2009).
51. Brown, N. G. et al. Mechanism of polyubiquitination by human anaphase-promoting complex: RING repurposing for ubiquitin chain assembly. *Mol. Cell* **56**, 246–260 (2014).
52. Weissmann, F. et al. biGbac enables rapid gene assembly for the expression of large multisubunit protein complexes. *Proc. Natl Acad. Sci. USA* **113**, E2564–E2569 (2016).
53. Hao, B., Oehlmann, S., Sowa, M. E., Harper, J. W. & Pavletich, N. P. Structure of a Fbw7-Skp1-cyclin E complex: multisite-phosphorylated substrate recognition by SCF ubiquitin ligases. *Mol. Cell* **26**, 131–143 (2007).
54. Schulman, B. A. et al. Insights into SCF ubiquitin ligases from the structure of the Skp1-Skp2 complex. *Nature* **408**, 381–386 (2000).
55. Walden, H. et al. The structure of the APPBP1-UBA3-NEDD8-ATP complex reveals the basis for selective ubiquitin-like protein activation by an E1. *Mol. Cell* **12**, 1427–1437 (2003).
56. Whitby, F. G., Xia, G., Pickart, C. M. & Hill, C. P. Crystal structure of the human ubiquitin-like protein NEDD8 and interactions with ubiquitin pathway enzymes. *J. Biol. Chem.* **273**, 34983–34991 (1998).
57. Koduri, V. et al. Peptidic degron for IMiD-induced degradation of heterologous proteins. *Proc. Natl Acad. Sci. USA* **116**, 2539–2544 (2019).
58. Starita, L. M. et al. Activity-enhancing mutations in an E3 ubiquitin ligase identified by high-throughput mutagenesis. *Proc. Natl Acad. Sci. USA* **110**, E1263–E1272 (2013).
59. Pierce, N. W., Kleiger, G., Shan, S. O. & Deshaies, R. J. Detection of sequential polyubiquitylation on a millisecond timescale. *Nature* **462**, 615–619 (2009).
60. Sievers, Q. L. et al. Defining the human C2H2 zinc finger degrome targeted by thalidomide analogs through CRBN. *Science* **362**, eaat0572 (2018).
61. Lu, G. et al. UBE2G1 governs the destruction of cereblon neomorphic substrates. *eLife* **7**, e40958 (2018).
62. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
63. Streich, F. C. Jr & Lima, C. D. Capturing a substrate in an activated RING E3/E2-SUMO complex. *Nature* **536**, 304–308 (2016).
64. Kastner, B. et al. GraFix: sample preparation for single-particle electron cryomicroscopy. *Nat. Methods* **5**, 53–55 (2008).
65. Palovcak, E. et al. A simple and robust procedure for preparing graphene-oxide cryo-EM grids. *J. Struct. Biol.* **204**, 80–84 (2018).
66. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
67. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
68. Hohn, M. et al. SPARX, a new environment for cryo-EM image processing. *J. Struct. Biol.* **157**, 47–55 (2007).
69. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
70. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
71. Afonine, P. V. et al. New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr. D* **74**, 814–840 (2018).
72. Duda, D. M. et al. Structure of a glomulin-RBX1-CUL1 complex: inhibition of a RING E3 ligase through masking of its E2-binding surface. *Mol. Cell* **47**, 371–382 (2012).
73. Yunus, A. A. & Lima, C. D. Lysine activation and functional analysis of E2-mediated conjugation in the SUMO pathway. *Nat. Struct. Mol. Biol.* **13**, 491–499 (2006).

Acknowledgements We thank D. Scott, J. Kellermann, J. Liwocha, S. Kostrhon, D. Horn-Ghetko, J. W. Harper, R. V. Farese Jr., H. Stark, D. Haselbach, S. Raunser, T. Raisch, S. Scheres, S. Uebel and S. Petter for assistance, reagents and helpful discussions; and M. Strauss, D. Bollschweiler, T. Schäfer and the cryo-EM facility at the Max Planck Institute of Biochemistry. This study was supported by the Max Planck Gesellschaft and by a grant of the European Commission (ERC Advanced Investigator Grant Nedd8Activate) to B.A.S. S.H. and G.K. were supported by a grant from the National Institutes of Health (R15GM117555-02).

Author contributions K.B., D.T.K., M.K., L.-M.N. and S.v.G. generated protein complexes and assayed quality for suitability for cryo-EM. D.T.K. conceived of, designed and generated stable proxies for ubiquitylation intermediates. K.B., D.T.K., S.H. and G.K. performed ubiquitylation assays. K.B. and J.R.P. collected, processed and refined cryo-EM data and built and refined structure. K.B., D.T.K., G.K. and B.A.S. prepared the manuscript with input from other authors. B.A.S. supervised the project.

Competing interests The authors declare no competing interests.

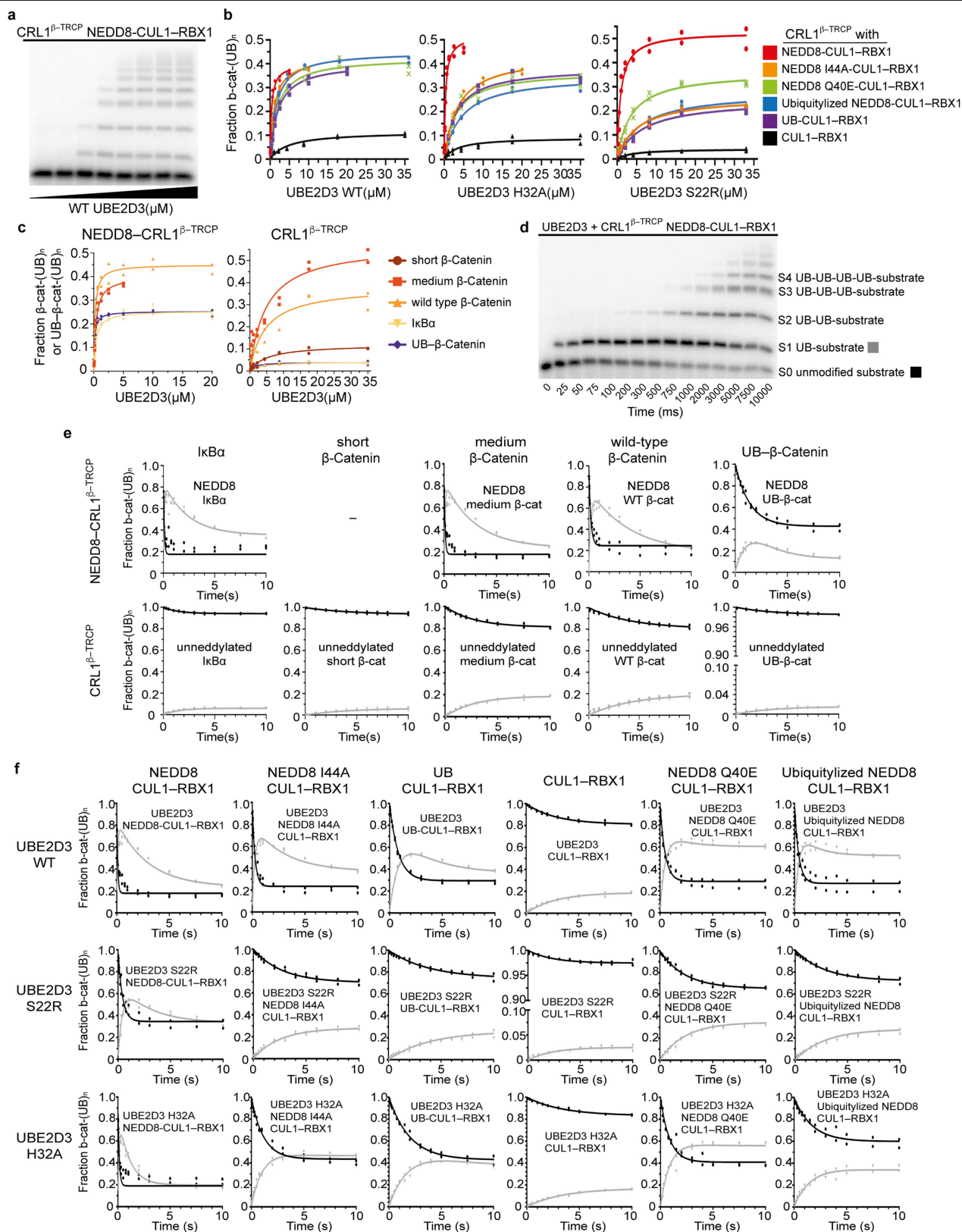
Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-2000-y>.

Correspondence and requests for materials should be addressed to B.A.S.

Peer review information Nature thanks David Barford, Ronald Hay and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

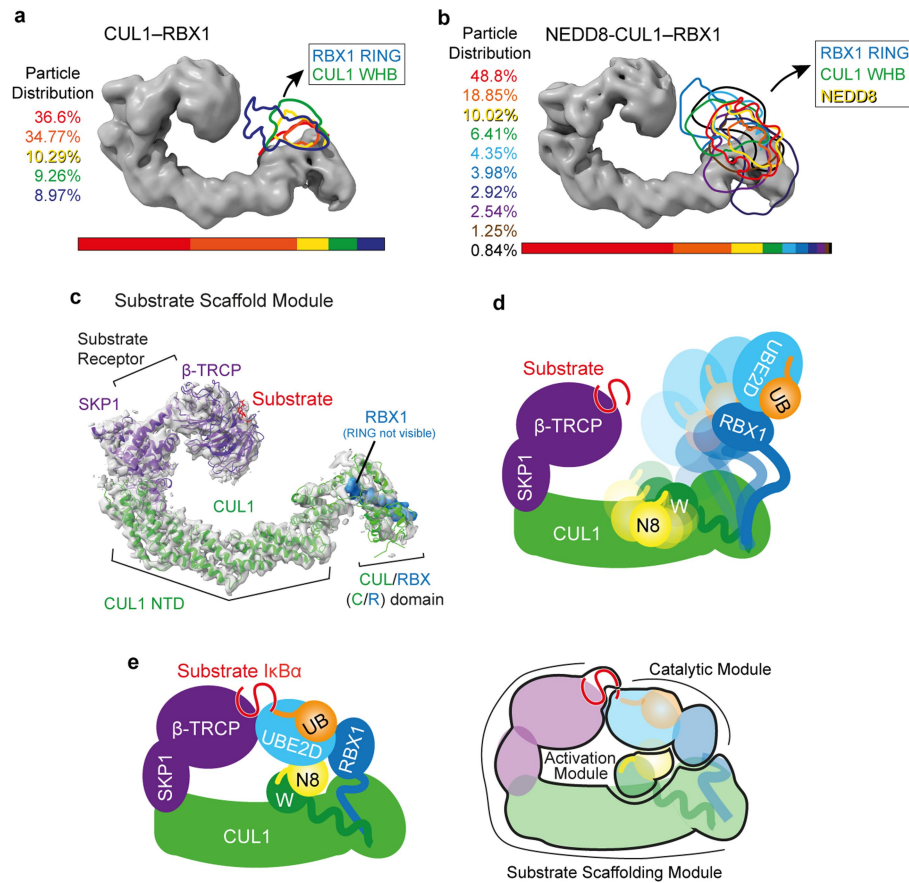
Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | See next page for caption.

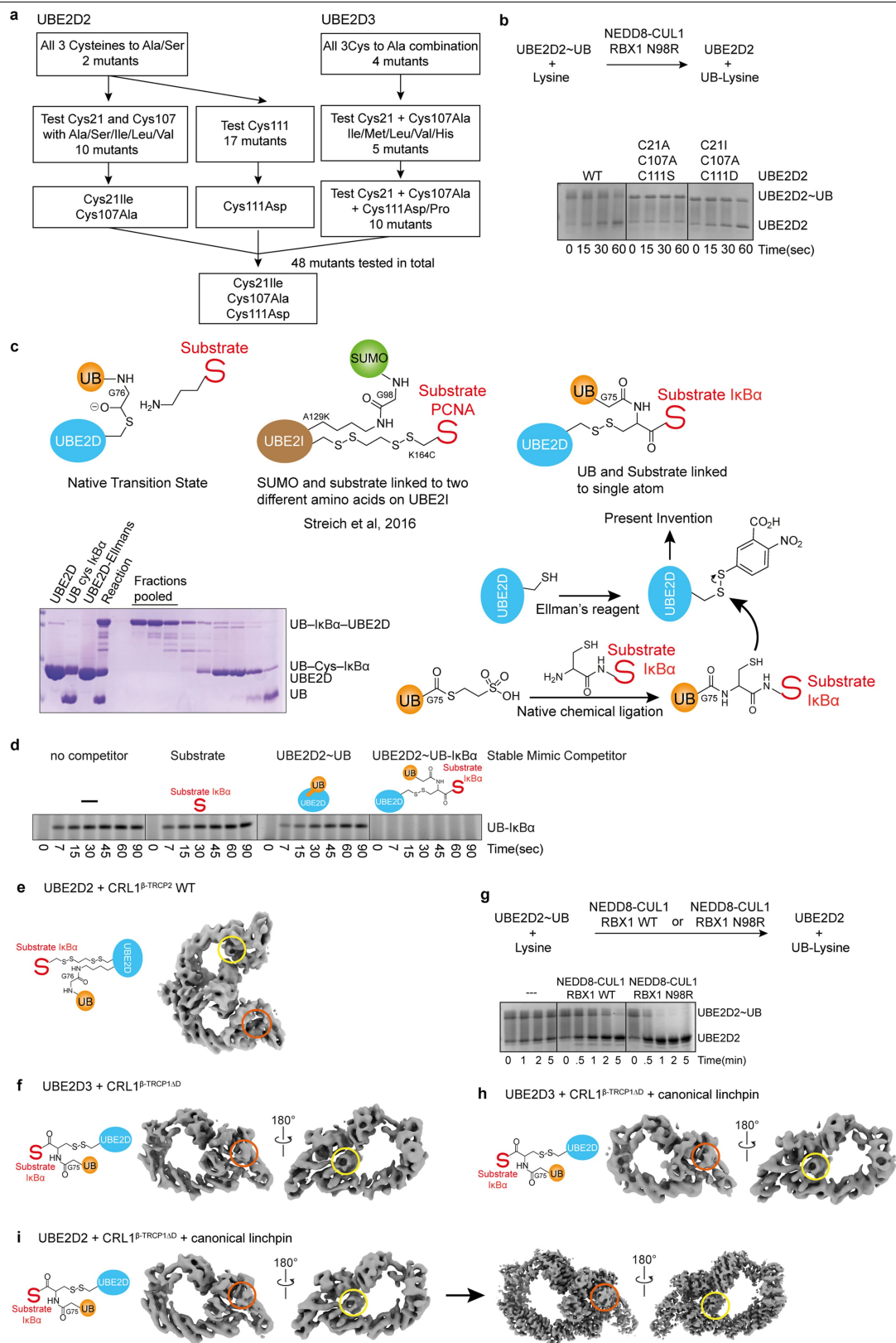
Extended Data Fig. 1 | Quantitative pre-steady-state enzyme kinetics of neddylated CRL1^{β-TRCP}- and UBE2D-dependent ubiquitylation. Gel images are representative of independent technical replicates ($n = 2$); the symbols on the graphs show the data from independent experiments ($n = 2$). **a**, Autoradiogram of SDS–PAGE gel showing products of ubiquitylation reactions under single-encounter conditions for the interaction of radiolabelled substrate (medium β -catenin substrate peptide derived from β -catenin) with neddylated CRL1^{β-TRCP}, titrating UBE2D3 (hereafter denoted UBE2D). Each lane represents a single ubiquitylation reaction that was used to estimate the fraction of peptide that had been converted into ubiquitylated products as a function of UBE2D concentration. **b**, Plots of the fraction of substrate that had been converted to ubiquitylated products against UBE2D concentration for ubiquitylation reactions containing either wild-type UBE2D (as shown in **a**), or the mutants UBE2D(S22R) or UBE2D(H32A). Various CRL1^{β-TRCP} complexes were assayed that contained either wild-type neddylated CRL1^{β-TRCP} (red), CRL1^{β-TRCP} complexes modified by NEDD8 variants containing I44A (orange), Q40E (green) or ‘ubiquitylizing’ (L2Q/K4F/E14T/D16E/G63K/G64E) substitutions (blue), or CRL1^{β-TRCP} modified by Ub(R72A) that is competent for ligation to CUL1 (purple) or unmodified CRL1^{β-TRCP} (CUL1 with the neddylation-site mutation K720R, black). Duplicate data points from independent experiments performed with identical samples are shown and were fit to the Michaelis–Menten model to estimate the K_m of UBE2D for CRL1^{β-TRCP} using nonlinear curve fitting (GraphPad Prism). **c**, Plots of the fraction of substrate that had been converted to ubiquitylated products against UBE2D concentration for ubiquitylation reactions with various substrate peptides: derived from I κ B α (but with a single acceptor Lys); derived from β -catenin; derived from β -catenin with different spacing between the phosphodegron motif and a potential acceptor Lys (a medium β -catenin substrate peptide with a nine-residue spacer between

the β -catenin phosphodegron and acceptor, matching the relative position of these moieties in I κ B α ; and a short β -catenin substrate in which the four residues between these moieties are too few to bridge the structurally observed gap between the substrate receptor and UBE2D-Ub active site); and a homogeneous ubiquitin linked- β -catenin generated by sortase-mediated transpeptidation wherein the only lysines are from ubiquitin. **d**, Autoradiogram of SDS–PAGE gel showing results from rapid quench-flow reactions under pre-steady-state single encounter conditions for the interaction of radiolabelled substrate (a medium β -catenin phosphopeptide) with CRL1^{β-TRCP}. The representative raw data are from a reaction using wild-type UBE2D and wild-type neddylated CRL1^{β-TRCP}, and show time-resolved conjugation of increasing numbers of individual ubiquitin molecules. S0, substrate with 0 ubiquitins; S1, substrate with one ubiquitin; S2, substrate with two ubiquitins; and so on. **e**, Plots comparing various substrate peptides described in **c**, showing disappearance of unmodified substrate (S0) with black circles, and the appearance of mono-ubiquitylated substrate (S1) with grey triangles, in rapid quench-flow reactions all performed as in **d** and under single-encounter conditions as in **a**. Duplicate data points from independent experiments performed with identical samples are shown. The data were fit to closed form equations (Mathematica) as previously described⁵⁹ to obtain both the rates for the transfer of the first ubiquitin to substrate (k_{obs}^{S0-S1}) and of the second ubiquitin to the singly ubiquitin-modified substrate (k_{obs}^{S1-S2}) as well as their associated standard error (Extended Data Table 1). **f**, Plots from experiments performed and analysed as described in **e**, except with radiolabelled medium β -catenin peptide substrate, CRL1^{β-TRCP} variants containing the indicated versions of CUL1–RBX1, and with either wild-type or indicated mutant versions of UBE2D.



Extended Data Fig. 2 | In CRL1 ^{β -TRCP}, the RING domain of RBX1 and the WHB domain of CUL1—without or with a covalently linked NEDD8—are dynamic in the absence of other factors and are harnessed in the catalytic architecture for substrate ubiquitylation with UBE2D. **a, Cryo-EM density corresponding to substrate-scaffolding regions of unneddylated CRL1 ^{β -TRCP} is shown in surface representation, with the density encompassing the RING domain of RBX1 and the WHB domain of CUL1 outlined for different 3D classes in different colours corresponding to the percentage of particles in that 3D class. **b**, As in **a** but with neddylated CRL1 ^{β -TRCP}, with its surfaces outlined in different classes encompassing the RING domain of RBX1, the WHB domain of CUL1, and covalently modified NEDD8. **c**, Refined cryo-EM density from CRL1 ^{β -TRCP} reveals the substrate-scaffolding module bridging the substrate recruited to substrate receptor β -TRCP with the intermolecular C/R domain, readily fitted with**

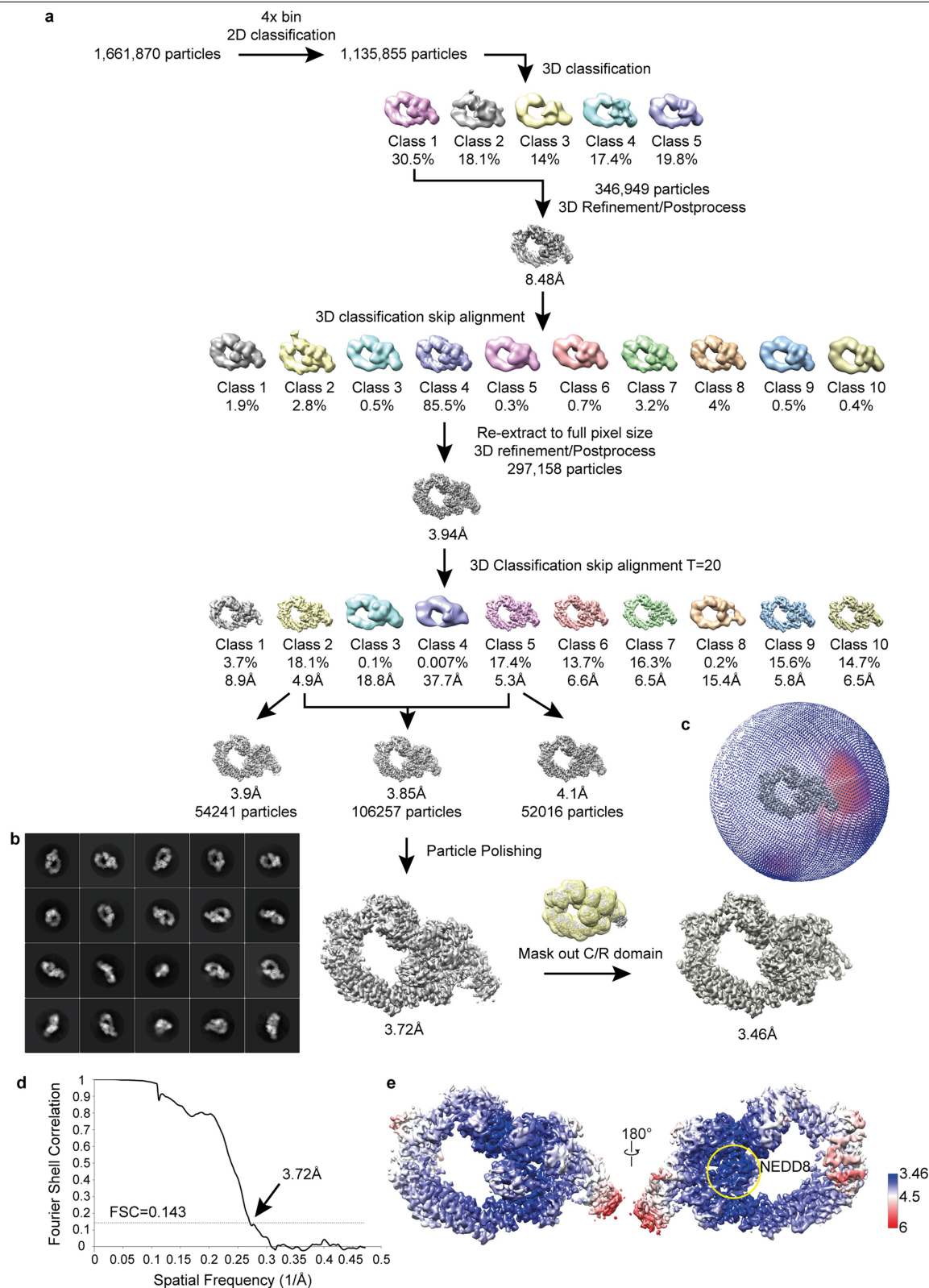
crystal structures of SKP1- β -TRCP (PDB: 1P22)²⁰ and the N-terminal domain of CUL1 and the C/R domain of CUL1-RBX1 (PDB: 1LDK)⁷. **d**, Schematic showing dynamics of NEDD8, its linked CUL1 WHB domain and the RING domain of RBX1 based on cryo-EM data in **b** for substrate-bound neddylated CRL1 ^{β -TRCP}, and model for varying locations of the RBX1 RING-bound UBE2D-Ub relative to the substrate awaiting ubiquitylation. **e**, Left, schematic of the catalytic architecture based on the cryo-EM data shown in Fig. 2, representing neddylated CRL1 ^{β -TRCP}-catalysed ubiquitin transfer from E2 UBE2D to an IkB α -derived substrate peptide. Right, semi-transparent version of the schematic, highlighting the three modules (substrate-scaffolding, catalytic and activation modules), their constituents and locations establishing the catalytic architecture for substrate priming by neddylated CRL1 ^{β -TRCP} and UBE2D.



Extended Data Fig. 3 | See next page for caption.

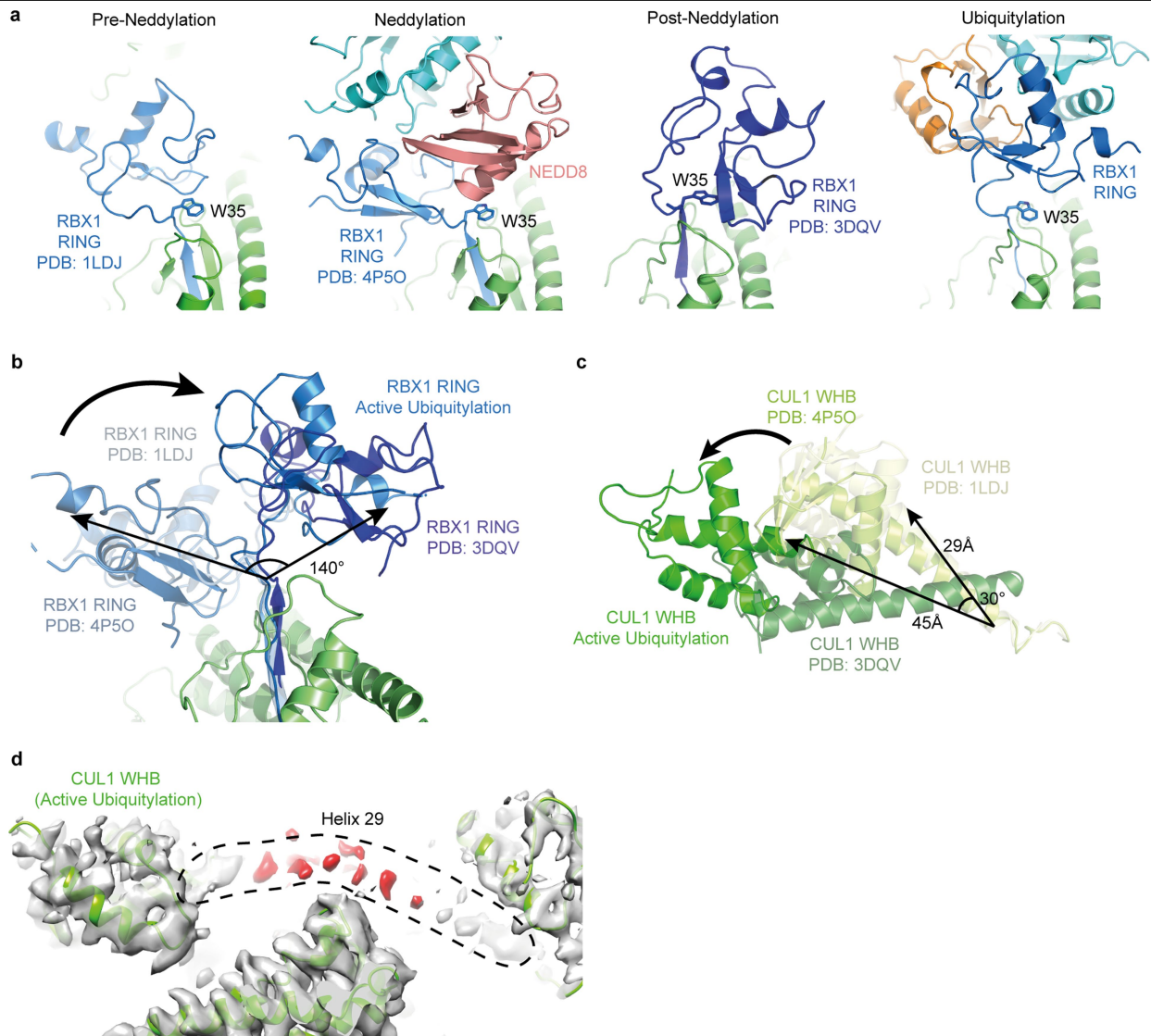
Extended Data Fig. 3 | Generation of a stable proxy for the UBE2D-Ub-substrate intermediate, and characterization in complexes with neddylated CRL1^{β-TRCP} by cryo-EM and biochemistry. Gel panels in this figure are representative of two independent experiments; $n = 2$. **a**, Our strategy for trapping a mimic of the transient neddylated CRL E2-Ub-substrate complex requires that the E2 UBE2D contain only a single cysteine at the active site. However, UBE2D contains three additional cysteines (Cys21, Cys107 and Cys111). Standard replacements of cysteine by serine or alanine severely compromised activity. On the basis of the structural locations of these cysteines, we presumed that their mutation hindered formation of the RING-activated, closed, active UBE2D-Ub conformation^{28–30}. We thus devised a systematic structure- and random-based approach to identify suitable replacements that qualitatively maintain wild-type levels of activity with neddylated CRLs. Structural analysis showed that Cys21 and Cys107 are in close proximity, such that mutation of both residues to alanine may generate a destabilizing cavity at this site. Combining UBE2D2(C107A) with Cys21 mutated to isoleucine, leucine or valine to compensate for the reduced hydrophobic volume led to the identification of C21I(C107A) as a suitable version for testing all other possible replacements for Cys111. A similar approach was taken for UBE2D3. A total of 48 different versions of UBE2D were tested to identify the UBE2D(C21I/C107A/C111D) mutant for chemical trapping at the remaining active site cysteine. **b**, Top, schematic of pulse-chase assay testing intrinsic activation of thioester-linked UBE2D-Ub intermediates. Although this is often tested by monitoring RING-dependent discharge of ubiquitin from UBE2D to free lysine, RBX1 RING-dependent activity is limited in this assay owing to sequence constraints imposed by the requirements for binding to partners other than UBE2D³⁹. Nonetheless, substrate-independent activation of UBE2D-Ub can be readily visualized using CUL1 complexed with a previously described hyperactive mutant RBX1(N98R)³⁹, and high enzyme and lysine concentrations. UBE2D-Ub generated in a pulse reaction was mixed with NEDD8-modified CUL1-RBX1 (shown here with the N98R mutant) and free lysine, and ubiquitin discharge was monitored over time by Coomassie-stained SDS-PAGE (as shown by the representative gel at the bottom) demonstrating that standard serine or alanine mutations of noncatalytic cysteines compromised activity (shown for the mutant C21A/C107A/C111S), whereas the optimized mutant (C21I/C107A/C111D) retains activity similar to that of the wild type. **c**, Overview of the generation of our stable proxy for the phosphorylated IκBα substrate intermediate linked at a single atom, and comparison to the previous method used to visualize noncanonical Lys sumoylation⁶³. **d**, Experiment validating our stable proxy for the UBE2D-Ub-phosphorylated

IκBα substrate intermediate linked at a single atom, based on the hypothesis that its simultaneous occupation of the binding sites for the UBE2D-Ub intermediate and substrate should result in more potent inhibition of a neddylated CRL1^{β-TRCP}-dependent substrate priming reaction compared to the individual constituents of the complex. **e**, Cryo-EM reconstruction of neddylated CRL1^{β-TRCP2} (with full-length, dimeric β-TRCP2) bound to a mimic of UBE2D2-Ub-IκBα generated by adapting the method used previously to visualize noncanonical lysine sumoylation⁶³. Ubiquitin is isopeptide-bonded to the substituted residue of a UBE2D(L119K) mutant, and a cysteine residue that replaces the acceptor in the substrate is disulfide-bonded to the catalytic cysteine of UBE2D2. This electron microscopy map visualizes the catalytic architecture of dimeric CRL1^{β-TRCP2} in which the dimerization domain agrees well with the previous crystal structure²⁷, and its linked NEDD8 (circled in yellow) is bound to the backside of UBE2D, but the donor ubiquitin (absent from the region circled in orange) was not visible—presumably owing to inadequacies of the method used to generate this mimic of the catalytic intermediate, in which the ubiquitin and substrate are not both simultaneously linked to the UBE2D catalytic cysteine. Variations between the two protomers of the dimer also exacerbated sample heterogeneity. **f**, Cryo-EM reconstruction of neddylated CRL1^{β-TRCP1ΔD} (with monomeric version of β-TRCP1, from residue 175 to the C terminus²⁰) bound to our newly developed proxy for the UBE2D3-Ub-IκBα intermediate. The phospho-IκBα peptide-substrate-bound β-TRCP-SKP1-CUL1-RBX1-NEDD8-UBE2D portion of this map superimposes with the map for the dimeric complex shown in **e**, but here the entire complex is visible—including both the NEDD8 (circled in yellow) and donor ubiquitin (circled in orange). **g**, To further increase cryo-EM sample homogeneity, we considered that the RBX1 RING sequence represents a compromise to meet requirements for its many different catalytic activities achieved with neddylation E2s, various ubiquitin carrying enzymes, and regulators including the inhibitor GLMN⁷². Therefore, we introduced a second RBX1 linchpin residue via mutation (N98R), which has previously been shown to improve neddylated CRL and UBE2D-dependent substrate priming at the expense of other RBX1-dependent functions (for example, with UBE2M and UBE2R2)³⁹. A Coomassie-stained SDS-PAGE gel from an assay for the intrinsic activity of UBE2D-Ub is shown, showing enhanced neddylated CRL-dependent activation of discharge to free lysine with the RBX1 N98R mutation. **h, i**, Cryo-EM reconstructions of neddylated CRL1^{β-TRCP1ΔD} with RBX1(N98R) bound to our newly developed proxies for the UBE2D3-Ub-IκBα and UBE2D2-Ub-IκBα intermediates, the latter of which was pursued for high-resolution electron microscopy (final reconstruction refined to 3.7 Å resolution, shown on right).



Extended Data Fig. 4 | Flow chart showing the stages of cryo-EM image processing. **a**, Cryo-EM image-processing flow chart. Ultimately, reconstruction of the data yielded a focused refinement at 3.46 Å resolution and a global refinement at 3.7 Å resolution that superimposes well with lower-resolution maps that were obtained during attempts to visualize substrate priming with neddylated wild-type dimeric CRL1^{β-TRCP}. **b**, Two-dimensional

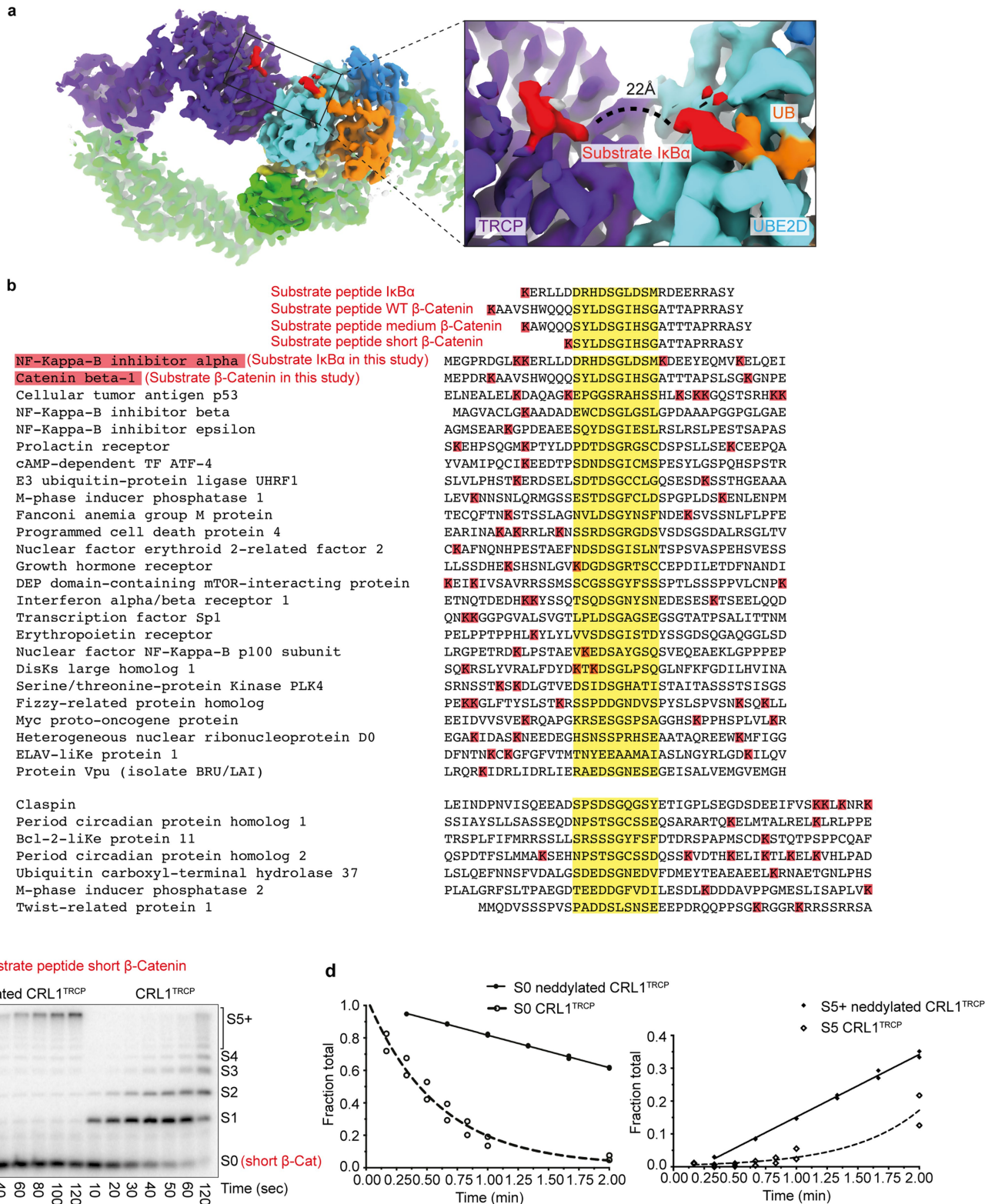
classes representing particles used for final reconstructions. **c**, Angular distribution of final reconstruction. **d**, Gold-standard Fourier shell correlation (FSC) curve showing overall resolution at 3.72 Å at an FSC of 0.143. **e**, Electron microscopy density map coloured by local resolution. NEDD8, circled in yellow, is the entity displaying the highest local resolution in the map.



Extended Data Fig. 5 | Extraordinary cullin-RING conformational changes in catalytic architecture juxtaposing the substrate and the active site of ubiquitylation. **a**, Side-by-side comparison of relative RING-domain locations in different CRL complexes after superposition of the C/R domains from the original CUL1-RBX1 structure (PDB: 1LDJ, 'pre-neddylation'—which data herein show is dynamic—although the crystal structure probably captured the conformation that enabled CAND1 binding and substrate receptor exchange)⁷, the structure representing the neddylation reaction (PDB: 4P50)³⁹, and a structure of a neddylated CUL5-RBX1 domain (PDB: 3DQV, labelled 'post-neddylation', which revealed the potential for conformational changes in the neddylated CUL WHB- and RBX1 RING-domains³, and data herein shows is dynamic), and the structure presented here showing how the neddylated CUL1 WHB domain and RBX1 RING domain are harnessed in a catalytic architecture for 'active ubiquitylation'. Trp35 of RBX1 is highlighted to show how it serves as a multifunctional platform for either the RING domain in different

orientations, or for the E2-linked NEDD8 during neddylation³⁹.

b, Superposition of the structures shown in **a**, highlighting different relative positions of the RING domain. **c**, Comparison of the relative locations of the CUL WHB domain in different structures after superimposing their C/R domains (not shown). **d**, Cryo-EM density from the neddylated CRL1^{B-TRCP}-UBE2D-Ub-substrate intermediate complex, showing patchiness of the region corresponding to CUL1 'helix-29'⁷. This CUL1 region connecting the C/R and WHB domains is visible only as patchy density, whereas in previous cullin crystals this forms the rod-like helix-29 continuing into the WHB domain⁷. It seems that helix-29 of CUL1 dissolves into a flexible tether, which rationalizes the previously observed proteolytic sensitivity of this region in a neddylated CUL1-RBX1 complex³, and enables the displacement and rotation required for placing the ensuing WHB domain and its linked NEDD8 at the centre of the ubiquitylation complex.



Extended Data Fig. 6 | See next page for caption.

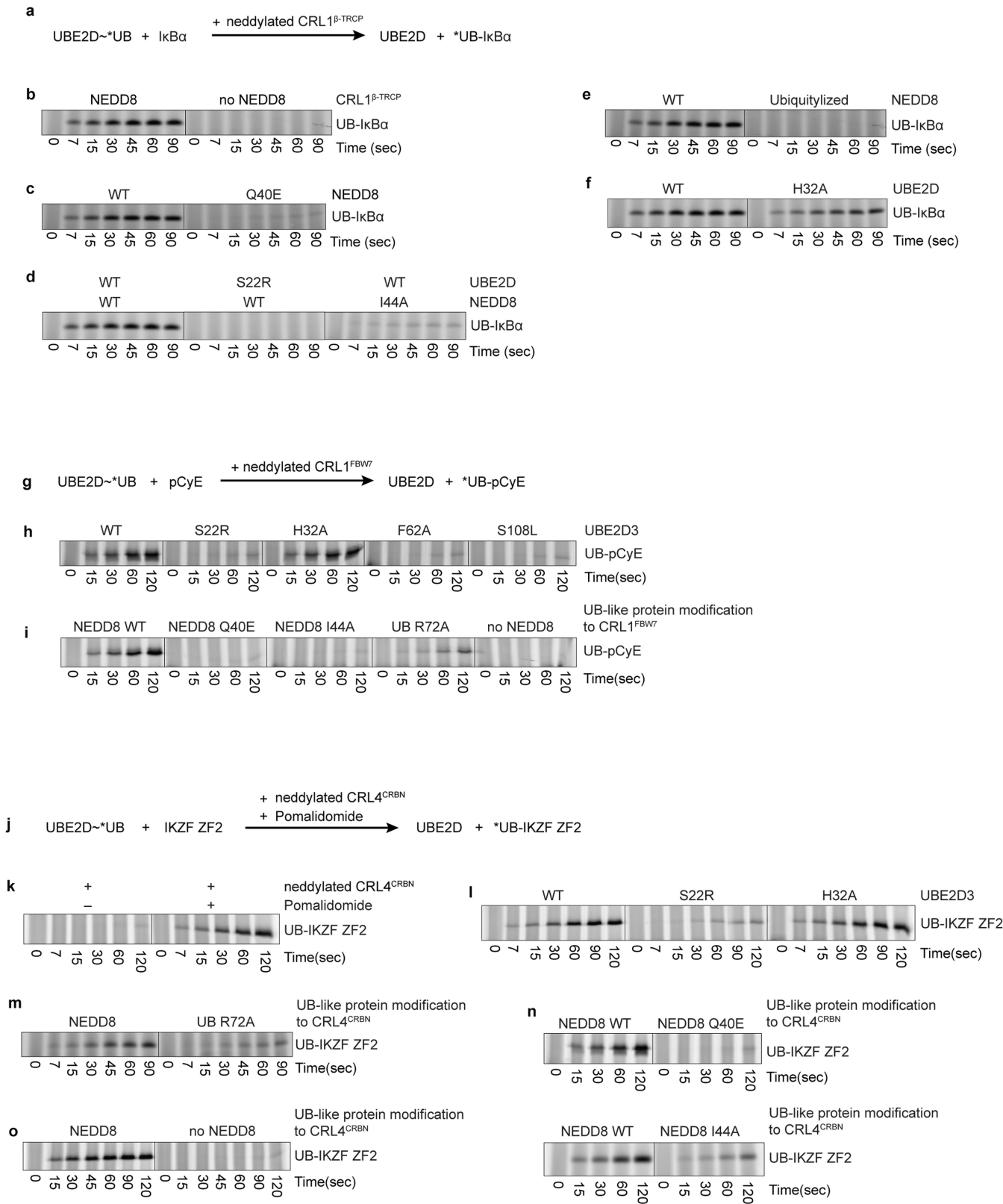
Extended Data Fig. 6 | Geometry between phosphodegron and acceptor in structure, substrates and ubiquitylation. **a**, Cryo-EM density highlighting the relative placement of the substrate degron and the UBE2D-Ub active site. The approximately 22 Å distance between the UBE2D-Ub active site and the phosphodegron of β -TRCP-bound substrate requires at least 6 intervening residues in a substrate. **b**, Alignments for several reported β -TRCP substrates³², highlighting the degron sequence (yellow) and nearby lysines (red). Also shown are sequences of peptide substrates with a single acceptor Lys that were used in kinetics analyses. The peptide sequences were derived from I κ B α , and from β -catenin with varying spacers between phosphodegron and acceptor Lys: wild-type β -catenin peptide, medium β -catenin peptide with lysine corresponding to I κ B α , and short β -catenin peptide with a lysine five residues upstream of the N-terminal phosphoserine in the degron, which would be too short to bridge the structurally observed distance between the

phosphodegron binding site on β -TRCP and the UBE2D catalytic cysteine in the ubiquitylation active site. **c**, Representative autoradiogram ($n = 2$) of SDS-PAGE gel showing products from indicated time points of ubiquitylation reactions under multiturnover conditions with either neddylated or unneddylated CRL1 ^{β -TRCP} and radiolabelled short β -catenin peptide substrate. The amount of short β -catenin peptide modified by neddylated CRL1 ^{β -TRCP} and UBE2D is too low in the single-encounter ubiquitylation reaction to enable quantification of kinetic parameters; however, product formation is apparent under multi-turnover conditions and shows that most products are heavily ubiquitylated. **d**, Plots fitting consumption of unmodified short β -catenin peptide substrate (S0) compared to formation of polyubiquitin chains with five or more ubiquitins (S5+) from reactions as in **c**. The symbols show the data from independent experiments ($n = 2$ technical replicates).

Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Interactions shaping the catalytic architecture of neddylated CRL1^{β-TRCP}-UBE2D-Ub-IkBα substrate intermediate. **a**, NEDD8 and the catalytic module from the structure representing the neddylated CRL1^{β-TRCP}-UBE2D-Ub-IkBα intermediate, highlighting distinctive interactions between NEDD8 (yellow) and donor ubiquitin (orange) with UBE2D. **b**, Catalytic module from the neddylated CRL1^{β-TRCP}-UBE2D-Ub-IkBα intermediate, highlighting the covalently linked proxy for the IkBα substrate's acceptor in the active site relative to a superimposed representative previous crystal structure of an isolated RING-UBE2D-Ub complex (grey, PDB: 4AP4)^{28,29}. In the inset, the density for the covalently linked proxy for IkBα substrate's acceptor is shown in red in the active site. The chemical trap superimposes with consensus acceptors visualized in active sites of sumoylation⁶³ and neddylation³⁹ intermediates, where aromatic side chains guide the lysine targets (blue and green, respectively)^{39,73}. However, the myriad substrates of UBE2D neither conform to a specific motif nor do they or UBE2D display specific side chains that guide lysine acceptors into the catalytic centre. Instead, in the neddylated CRL1^{β-TRCP}-UBE2D-Ub-substrate complex, density from backbone atoms preceding the chemical proxy for the acceptor lysine corresponds to the aromatic guides in sumoylation and neddylation intermediates. **c**, Overview of assays for activation of intrinsic reactivity of the UBE2D-Ub intermediate. Top, schematic of pulse-chase assay for testing the effects of UBE2D mutations on activation, monitoring UBE2D-Ub discharge to free lysine activated by neddylated CUL-RBX1 compared to unneddylated or RING-like UBE4B controls. Bottom, sites of mutations shown as spheres on the structure of UBE2D from the cryo-EM structure of neddylated CRL1^{β-TRCP}-UBE2D-Ub-substrate complex. The colours of spheres reflect both the locations and the effects on UBE2D-Ub discharge to free lysine. Sites of mutations with marginal or no effect are shown in cyan, whereas those with major effects are otherwise

coloured. Mutations that cause major defects map to the RBX1 RING-binding site (blue), the interaction surface with the donor ubiquitin (orange), and the interaction surface with NEDD8 (yellow). **d**, Representative Coomassie-stained SDS-PAGE gels (of two independent experiments) shown for reactions monitoring substrate-independent discharge of UBE2D-Ub to free lysine, in the presence of CUL1-RBX1(N98R) that was either neddylated or unneddylated (K720R), with either wild-type or the indicated UBE2D3 mutants at binding sites for backside-bound NEDD8(S22R), the RBX1 RING(F62A), and the covalently linked donor ubiquitin in the closed conformation (S108L). **e**, As in **d**, except testing the effect of NEDD8(Q40E), which would disrupt the activation module. **f**, Reactions performed as in **d**, except with indicated variants of UBE2D2, in reactions with CUL1-RBX1(N98R) that was either neddylated or unneddylated (K720R), or with the optimized RING-like U-box domain from UBE4B as a reference⁵⁸. For mutations reporting on the catalytic conformation (G24K, T36K, M38K, A96K and D112K), representative gels are shown for two experiments. All other experiments were performed once. **g**, Comparison of β1/β2-loop conformations after superimposing the indicated structures of NEDD8 and ubiquitin. The comparison suggests that whereas NEDD8 and ubiquitin can adopt both loop-in and loop-out conformations, donors linked to E2 active sites in RING-activated complexes adopt the loop-in conformation, and those bound to the UBE2D backside adopt loop-out conformations. **h**, Only a loop-out conformation is compatible with the neddylated CRL activation module structure, because loop-in conformation from the structures shown in **g** would prevent noncovalent interactions with the WHB domain of CUL1 (green). **i**, Only a loop-out conformation is compatible for the CUL1-linked NEDD8 to bind the catalytic module, because loop-in conformations in the structures shown in **g** would prevent noncovalent interactions with the backside of UBE2D (cyan).



Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Qualitative validation of the mechanistic principles that underlie substrate priming by neddylated CRLs and UBE2D. Gel panels in this figure are representative of two independent experiments; $n = 2$ technical replicates. **a**, Schematic of a qualitative substrate priming assay for testing effects of mutations in neddylated CRL1^{FBW7} or UBE2D on substrate priming, monitoring fluorescent ubiquitin transfer from UBE2D3 to the phosphorylated IκBα substrate. **b**, Scan of SDS-PAGE detecting fluorescent ubiquitin transferred to the IκBα-derived substrate in a qualitative assay for NEDD8 activation of substrate priming. **c**, As in **b**, showing the effect on substrate priming of disrupting the activation module with the Q40E mutation of NEDD8. **d**, As in **b**, showing the effect on substrate priming of disrupting interactions between the activation and catalytic modules with the NEDD8(I44A) or UBE2D(S22R) mutants. **e**, As in **b**, showing the effect on substrate priming of disrupting interactions between the activation and substrate-scaffolding modules, though CUL1 modification by a 'ubiquitylized' NEDD8 mutant with six residues swapped for their ubiquitin counterparts (L2Q/K4F/E14T/D16E/G63K/G64E). **f**, As in **b**, showing the effect of the H32A in UBE2D at the interface between the catalytic and substrate-scaffolding modules. **g**, Scheme of pulse-chase assay for testing the effects of mutations in

neddylated CRL1^{FBW7} or UBE2D on substrate priming. The assay monitors the transfer of fluorescent ubiquitin from UBE2D to peptide substrate derived from phosphorylated cyclin E (pCyE). **h**, Fluorescent scan detecting ubiquitin transferred to the pCyE substrate by neddylated CRL1^{FBW7} and the indicated mutants of UBE2D. **i**, Fluorescent scan detecting ubiquitin transferred to the pCyE substrate by UBE2D and indicated variants of neddylated (or ubiquitylated) CRL1^{FBW7}. Experiment with unneddylated CRL1^{FBW7} used the K720R variant of CUL1 to prevent artefactual ubiquitylation. **j**, Scheme of pulse-chase assay for testing effects of mutations in neddylated CRL4^{CRBN} or UBE2D on substrate priming, monitoring fluorescent ubiquitin transfer from UBE2D to the IKZF1/3 ZF2 substrate in the presence of the immunomodulatory drug pomalidomide. **k**, Fluorescent scan of assay validation, showing dependence on pomalidomide. **l**, Fluorescent scan detecting ubiquitin transferred to the IKZF substrate by CRL4^{CRBN}, pomalidomide and the indicated variants of UBE2D. **m-o**, Fluorescent scan detecting ubiquitin transferred to the IKZF substrate by UBE2D and the indicated variants of neddylated (or ubiquitylated) CRL4^{CRBN} with pomalidomide. Experiments with unneddylated CRL4^{CRBN} used the K705R variant of CUL4A to prevent artefactual ubiquitylation.

Extended Data Table 1 | Estimates of K_m and k_{obs} for substrate ubiquitylation

Extended Data Table 1. Estimates of K_m and k_{obs} for Substrate Ubiquitylation							
Substrate	E2	CRL1 β -TRCP	K_m (10^{-9} M)	k_{obs}^{S0-S1} (sec^{-1})	k_{obs}^{S1-S2} (sec^{-1})	k_{obs}^{S0-S1}/K_m ($\text{M}^{-1}\text{sec}^{-1}$)	Fold change
I κ B α	WT UBE2D	NEDD8-CUL1-RBX1	408 \pm 57	12.9 \pm 1.7	0.24 \pm 0.04	3.2 $\times 10^7$	-
I κ B α	WT UBE2D	CUL1-RBX1	4014 \pm 956	0.05 \pm 0.002	-	1.2 $\times 10^4$	2667
WT β -cat	WT UBE2D	NEDD8-CUL1-RBX1	214 \pm 30	4.8 \pm 0.46	0.19 \pm 0.02	2.2 $\times 10^7$	-
WT β -cat	WT UBE2D	CUL1-RBX1	4638 \pm 138	0.05 \pm 0.005	-	1.1 $\times 10^4$	2000
medium β -cat	WT UBE2D	NEDD8-CUL1-RBX1	372 \pm 48	11 \pm 1.00	0.24 \pm 0.04	3.0 $\times 10^7$	-
medium β -cat	WT UBE2D	CUL1-RBX1	4875 \pm 1328	0.08 \pm 0.005	-	1.6 $\times 10^4$	1875
short β -cat	WT UBE2D	CUL1-RBX1	5207 \pm 1051	0.02 \pm 0.0004	-	3.8 $\times 10^3$	-
medium β -cat	WT UBE2D	NEDD8 I44A-CUL1-RBX1	1087 \pm 69	3.60 \pm 0.27	0.17 \pm 0.03	3.3 $\times 10^6$	9.1
medium β -cat	WT UBE2D	UB-CUL1-RBX1	2056 \pm 170	1.0 \pm 0.03	0.15 \pm 0.01	4.9 $\times 10^5$	61.2
medium β -cat	WT UBE2D	NEDD8 Q40E-CUL1-RBX1	1941 \pm 226	1.77 \pm 0.13	0.07 \pm 0.01	9.1 $\times 10^5$	33.0
medium β -cat	WT UBE2D	UBylized NEDD8-CUL1-RBX1	1651 \pm 137	2.20 \pm 0.15	0.17 \pm 0.04	1.3 $\times 10^6$	23.1
medium β -cat	S22R UBE2D	NEDD8-CUL1-RBX1	913 \pm 135	1.89 \pm 0.14	0.17 \pm 0.03	2.1 $\times 10^6$	14.2
medium β -cat	S22R UBE2D	NEDD8 I44A-CUL1-RBX1	5190 \pm 604	0.13 \pm 0.01	-	2.5 $\times 10^4$	1200
medium β -cat	S22R UBE2D	UB-CUL1-RBX1	6492 \pm 872	0.08 \pm 0.005	-	1.2 $\times 10^4$	2500
medium β -cat	S22R UBE2D	CUL1-RBX1	4007 \pm 754	0.013 \pm 0.002	-	3.3 $\times 10^3$	9091
medium β -cat	S22R UBE2D	NEDD8 Q40E-CUL1-RBX1	3581 \pm 367	0.17 \pm 0.005	-	4.7 $\times 10^4$	638.3
medium β -cat	S22R UBE2D	UBylized NEDD8-CUL1-RBX1	6105 \pm 1030	0.12 \pm 0.007	-	2.0 $\times 10^4$	1500
medium β -cat	H32A UBE2D	NEDD8-CUL1-RBX1	347 \pm 53	8.15 \pm 1.04	0.73 \pm 0.09	2.3 $\times 10^7$	1.3
medium β -cat	H32A UBE2D	NEDD8 I44A-CUL1-RBX1	3431 \pm 332	0.58 \pm 0.04	0.12 \pm 0.02	1.7 $\times 10^5$	176.5
medium β -cat	H32A UBE2D	UB-CUL1-RBX1	2999 \pm 301	0.34 \pm 0.01	0.10 \pm 0.02	1.1 $\times 10^5$	272.7
medium β -cat	H32A UBE2D	CUL1-RBX1	3305 \pm 1119	0.06 \pm 0.003	-	1.8 $\times 10^4$	1666.7
medium β -cat	H32A UBE2D	NEDD8 Q40E-CUL1-RBX1	2923 \pm 459	0.81 \pm 0.06	-	2.8 $\times 10^5$	107.1
medium β -cat	H32A UBE2D	UBylized NEDD8-CUL1-RBX1	4303 \pm 310	0.28 \pm 0.02	-	6.5 $\times 10^4$	461.5
UB- β -cat	WT UBE2D	NEDD8-CUL1-RBX1	148 \pm 15	-	0.48 \pm 0.02	-	-
UB- β -cat	WT UBE2D	CUL1-RBX1	1503 \pm 439	-	0.005 \pm 0.0004	-	-

S0 refers to unmodified substrate, S1 to substrate modified by a single ubiquitin and S2 to substrate modified with two ubiquitins. S0-S1 refers to the transition from unmodified substrate to Ub-substrate, and S1-S2 to the transition from Ub-substrate to Ub-Ub-substrate. Values for K_m are the best-fit values derived from nonlinear regression in Prism, and value for k_{obs} are the best-fit values derived from nonlinear regression in Mathematica. The measure of error is the standard error of the mean as determined by Prism and Mathematica, respectively, from experiments and curve fits such as those shown in Extended Data Fig. 1 ($n = 2$).

Extended Data Table 2 | Cryo-EM data collection, refinement and validation statistics

Cryo-EM data collection, refinement and validation statistics							
IkBα-UB~UBE2D crosslink (C21H C107A C111D) Substrate Receptor	UBE2D2	UBE2D3	UBE2D3	UBE2D2	UBE2D2	none	none
	2x2way XL	3way XL	3way XL	3way XL	3way XL		
	β-TRCP2	β-TRCP1	β-TRCP1	β-TRCP1	β-TRCP1	β-TRCP1	β-TRCP1
		175-C	175-C	175-C	175-C	175-C	175-C
	WT	WT	N98R	N98R	N98R	WT	WT
RBX1	yes	yes	yes	yes	yes	no	yes
NEDD8	WT	ΔΔ	ΔΔ	ΔΔ	ΔΔ	WT	WT
SKP1	EMD-10578	EMD-10579	EMD-10580	EMD-10581	EMD-10585	EMD-10582	EMD-10583
PDB 6TTU							
Data collection and processing							
Microscope	Krios	Arctica	Arctica	Arctica	Krios	Glacios	Glacios
Magnification	105,000	92,000	73,000	73,000	130,000	36,000	22,000
Voltage (kV)	300	200	200	200	300	200	200
Electron exposure (e-/Å ²)	56	61.3	60.8	70	70.2	60	59
Defocus range (μm)	-1.2 ~ -3.6	-1.5 ~ -3.5	-1.5 ~ -3.5	-1.5 ~ -3.5	-1.2 ~ -3.3	-1.2 ~ -3.3	-1.2 ~ -3.3
Pixel size (Å)	1.34	1.612	1.997	1.997	1.06	1.181	1.885
Symmetry imposed	C2	C1	C1	C1	C1	C1	C1
Initial particle images (no.)	2,575,161	464,344	601,121	459,011	1,661,870	2,051,804	1,666,293
Final particle images (no.)	33,738	47,246	107,311	40,835	106,257	262,116	349,803
Map resolution (Å)	9.3	8.6	9.4	8.4	3.72	4.64	6.7
FSC threshold	(0.143)	(0.143)	(0.143)	(0.143)	(0.143)	(0.143)	(0.143)
Map resolution range (Å)	-	-	-	-	3.46 ~ 6.0	-	-
Refinement							
Initial model used (PDB code)					1LDJ 6M90 4P5O 4V3L		
Model resolution (Å)					3.7		
FSC threshold					(0.143)		
Map sharpening <i>B</i> factor (Å ²)	-578.9	-1159	-1272	-983.5	-94.2	-199	-338
Model composition							
Non-hydrogen atoms					13028		
Protein residues					1616		
Ligands					3(ZN)		
<i>B</i> factors (Å ²)							
Protein					88.6		
Ligand					203.8		
R.m.s. deviations							
Bond lengths (Å)					0.006		
Bond angles (°)					0.850		
Validation							
MolProbity score					2.23		
Clashscore					13.09		
Poor rotamers (%)					0.07		
Ramachandran plot							
Favored (%)					91		
Allowed (%)					9		
Disallowed (%)					0		

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | | |
|-------------------------------------|---|
| n/a | Confirmed |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Gel imaging: Amersham Imager 600, Amersham Typhoon, Typhoon 9410 Imager; Cryo-EM: SerialEM, FEI EPU

Data analysis

Gel Quantification and kinetics: ImageQuant v5.2, Mathematica v10.0.0, GraphPad Prism v8.2.1; Cryo-EM: RELION v3.0, SPARX v2.22, Gautomatch v0.56, CTFFIND v4.1.9; Structure Visualization: Chimera v1.11.2, ChimeraX v0.9, PyMol v2.3.3; Model Building: COOT v0.89, Phenix.refine v1.15.2

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Cryo-EM maps will be available from the Electron Microscopy Data Bank, and the model coordinates will be available from Protein Data Bank upon publication.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	All experiments we performed at least twice.
Data exclusions	No data were excluded.
Replication	All experiments we performed at least twice, with numerous controls. All attempts at replication were successful.
Randomization	No grouped samples.
Blinding	No grouped samples.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Insights into the assembly and activation of the microtubule nucleator γ -TuRC

<https://doi.org/10.1038/s41586-019-1896-6>

Received: 20 August 2019

Accepted: 10 December 2019

Published online: 19 December 2019

Peng Liu^{1,6}, Erik Zupa^{1,6}, Annett Neuner¹, Anna Böhler¹, Justus Loerke², Dirk Flemming³, Thomas Ruppert¹, Till Rudack⁴, Christoph Peter⁵, Christian Spahn², Oliver J. Gruss⁵, Stefan Pfeffer^{1*} & Elmar Schiebel^{1*}

Microtubules are dynamic polymers of α - and β -tubulin and have crucial roles in cell signalling, cell migration, intracellular transport and chromosome segregation¹. They assemble de novo from $\alpha\beta$ -tubulin dimers in an essential process termed microtubule nucleation. Complexes that contain the protein γ -tubulin serve as structural templates for the microtubule nucleation reaction². In vertebrates, microtubules are nucleated by the 2.2-megadalton γ -tubulin ring complex (γ -TuRC), which comprises γ -tubulin, five related γ -tubulin complex proteins (GCP2–GCP6) and additional factors³. GCP6 is unique among the GCP proteins because it carries an extended insertion domain of unknown function. Our understanding of microtubule formation in cells and tissues is limited by a lack of high-resolution structural information on the γ -TuRC. Here we present the cryo-electron microscopy structure of γ -TuRC from *Xenopus laevis* at 4.8 Å global resolution, and identify a 14-spoked arrangement of GCP proteins and γ -tubulins in a partially flexible open left-handed spiral with a uniform sequence of GCP variants. By forming specific interactions with other GCP proteins, the GCP6-specific insertion domain acts as a scaffold for the assembly of the γ -TuRC. Unexpectedly, we identify actin as a bona fide structural component of the γ -TuRC with functional relevance in microtubule nucleation. The spiral geometry of γ -TuRC is suboptimal for microtubule nucleation and a controlled conformational rearrangement of the γ -TuRC is required for its activation. Collectively, our cryo-electron microscopy reconstructions provide detailed insights into the molecular organization, assembly and activation mechanism of vertebrate γ -TuRC, and will serve as a framework for the mechanistic understanding of fundamental biological processes associated with microtubule nucleation, such as meiotic and mitotic spindle formation and centriole biogenesis⁴.

To understand the structural basis of microtubule (MT) nucleation in vertebrates, we purified γ -TuRCs from *X. laevis* meiotic egg extract by affinity chromatography, and confirmed the integrity of the complexes by MT nucleation analysis, immunoblotting, sucrose gradient density sedimentation and negative-stain electron microscopy (Extended Data Fig. 1). To gain insights into the relative abundance of GCP variants in the complex, we analysed the purified γ -TuRC by label-free quantification (LFQ) mass spectrometry and determined the stoichiometry of the components normalized to a 14-spoked γ -TuRC. The purified γ -TuRC contains five copies of GCP2, five copies of GCP3, two or three copies of GCP4 and one copy of GCP5 and GCP6, resulting in a 5:5:2/3:1:1 stoichiometry (Figs. 1a, 2a).

Molecular architecture of the γ -TuRC

Using cryo-electron microscopy (cryo-EM) single-particle analysis, we obtained reconstructions of the purified γ -TuRC (Extended Data

Fig. 2a–e). Local resolution assessment for initial reconstructions indicated conformational flexibility (Extended Data Fig. 2f), which we compensated for computationally by splitting the γ -TuRC into segments that were refined independently (Extended Data Fig. 2e). This approach resulted in a final reconstruction at 4.8 Å global resolution, ranging from 4.5 Å to 6 Å locally (Extended Data Fig. 2g, h). The reconstruction (Fig. 1a) indicates that the γ -TuRC consists of 14 structurally similar spokes arranged in an open left-handed spiral with a diameter of approximately 32 nm and a height of around 25 nm. Each spoke consists of one GCP and one γ -tubulin bound to the conserved GCP GRIP2 domain (Fig. 1b). Resolved features distinguish the structurally closely related GCPs in the different positions and indicate that the GCP variants assemble in a defined order. Notably, the N-terminal extensions present in all GCPs except GCP4 (Extended Data Fig. 1a) were mostly not resolved, which suggests a high degree of conformational flexibility for these regions.

¹Zentrum für Molekulare Biologie, Universität Heidelberg, DKFZ-ZMBH Allianz, Heidelberg, Germany. ²Charité - Universitätsmedizin Berlin, Institute for Medical Physics and Biophysics, Berlin, Germany. ³Heidelberg University Biochemistry Center, Heidelberg, Germany. ⁴Department of Biophysics, Ruhr-University Bochum, Bochum, Germany. ⁵Institut für Genetik Universität Bonn, Bonn, Germany. ⁶These authors contributed equally: Peng Liu, Erik Zupa. *e-mail: s.pfeffer@zmbh.uni-heidelberg.de; e.schiebel@zmbh.uni-heidelberg.de

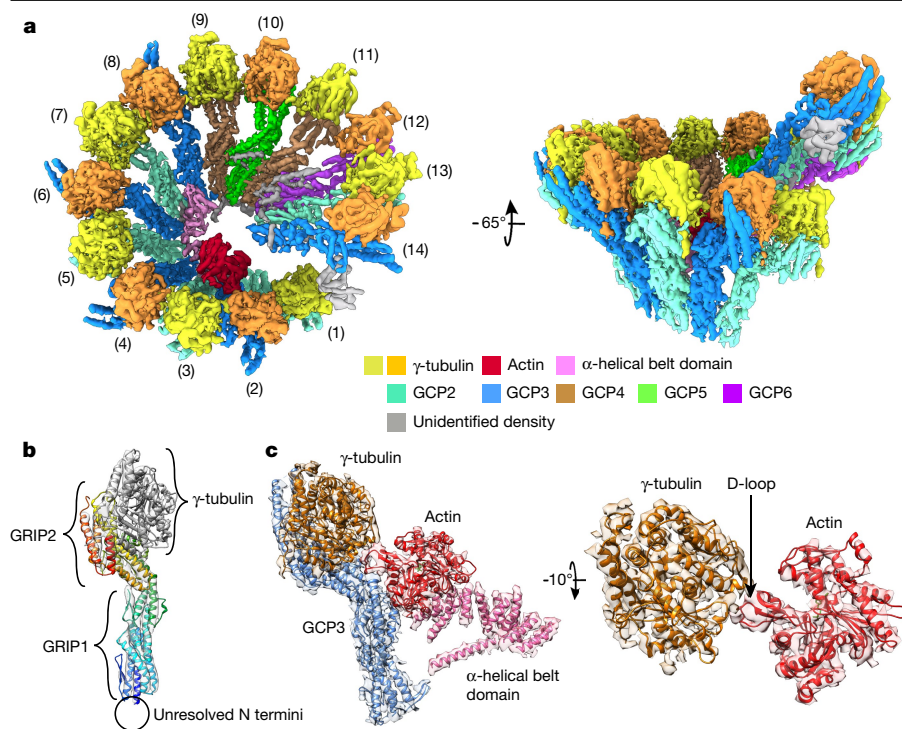


Fig. 1 | Cryo-EM structure of the *X. laevis* γ-TuRC. **a, Reconstruction of the γ-TuRC, filtered according to local resolution. For each spoke, the consecutive numbering is given. **b**, General layout of a γ-TuRC spoke. Atomic models for γ-tubulin (grey) and GCP2 (rainbow-coloured from the N (blue) to the C (red) terminus) superposed to the cryo-EM density. **c**, Detailed view on molecular components involved in binding of actin.**

For an initial investigation of the γ-TuRC architecture, we structurally grouped the 14 GCP-γ-tubulin spokes by computing pairwise cross-correlation coefficients between density segments for each spoke (Fig. 2b, Extended Data Fig. 3a) and by computing the pairwise root mean square deviation (r.m.s.d.) values of atomic models for human GCP4 (Protein Data Bank (PDB) code 3RIP) and γ-tubulin (PDB code 1Z5W) docked into all 14 spokes (Fig. 2b, Extended Data Fig. 3b). To account for the different conformational states of GCPs, the GCP4 model was split into three individual segments (Methods), which were then docked independently. Both approaches clustered the spokes into 5 groups with 5:5:2:1:1 stoichiometry, namely: (i) spokes 1, 3, 5, 7 and 13; (ii) spokes 2, 4, 6, 8 and 14; (iii) spokes 9 and 11; (iv) spoke 10; and (v) spoke 12. In combination with LFQ mass spectrometry (Fig. 2a), this indicates that the most abundant GCP groups (i) and (ii) correspond to GCP2 and GCP3. Most prominently,

GCPs in these two groups were structurally distinguished by the length of two α-helices in the GRIP2 domain (Fig. 2c, Extended Data Fig. 3c). These two helices were predicted to be considerably longer in GCP3 than in all other GCP variants (Extended Data Fig. 4a), which allows unambiguous assignment of group (ii) to GCP3 and conversely indicates that GCPs in group (i) correspond to GCP2. This could be confirmed by a GCP2-specific extended loop (Extended Data Fig. 4b) between two GRIP2 β-strands visible exclusively in the density for group (i) spokes (Fig. 2c, Extended Data Fig. 3c). Structural clustering in conjunction with LFQ mass spectrometry also suggested that group (iii) with spokes at positions 9 and 11 corresponded to GCP4. GCP4 is the only GCP variant without an N-terminal extension (Extended Data Fig. 1a), and only group (iii) spokes had no continuous density extending from the very N-terminal helix (Fig. 2c, Extended Data Fig. 3d). Although the two remaining spokes (spokes

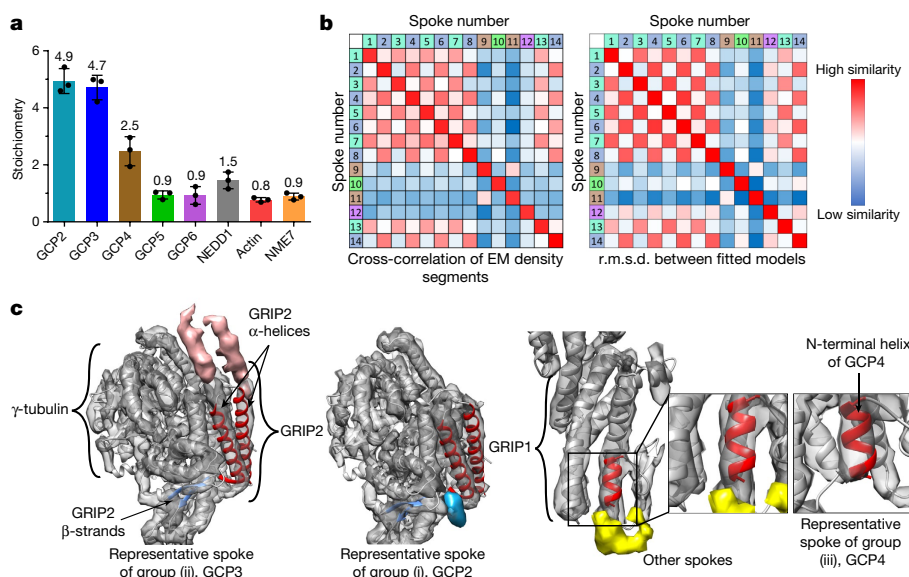


Fig. 2 | Structural clustering of γ-TuRC spokes. **a**, Relative abundance of γ-TuRC components as determined by LFQ mass spectrometry. Total number of GCP proteins is normalized to 14, and the abundance of each protein is calculated accordingly. $n = 3$ biologically independent experiments; data are mean \pm s.d. **b**, Left, pairwise cross-correlation coefficients between density segments. Extended Data Figure 3a provides correlation values. Right, pairwise r.m.s.d. values between Cα atoms of rigid-body-fitted atomic models (Methods) representing the individual spokes. Extended Data Figure 3b provides r.m.s.d. values. **c**, Cluster-specific structural features. Left, GCP3-specific extended C-terminal α-helices (red model and density) are unique for group (ii), and a GCP2-specific extended loop between the GRIP2 β-strands (blue model and density) is present only in group (i). Right, only group (iii) spokes are devoid of continuous density (yellow) connecting to the N-terminal helix (red) of the fitted GCP4 model. See Extended Data Fig. 3c, d for features for all spokes.

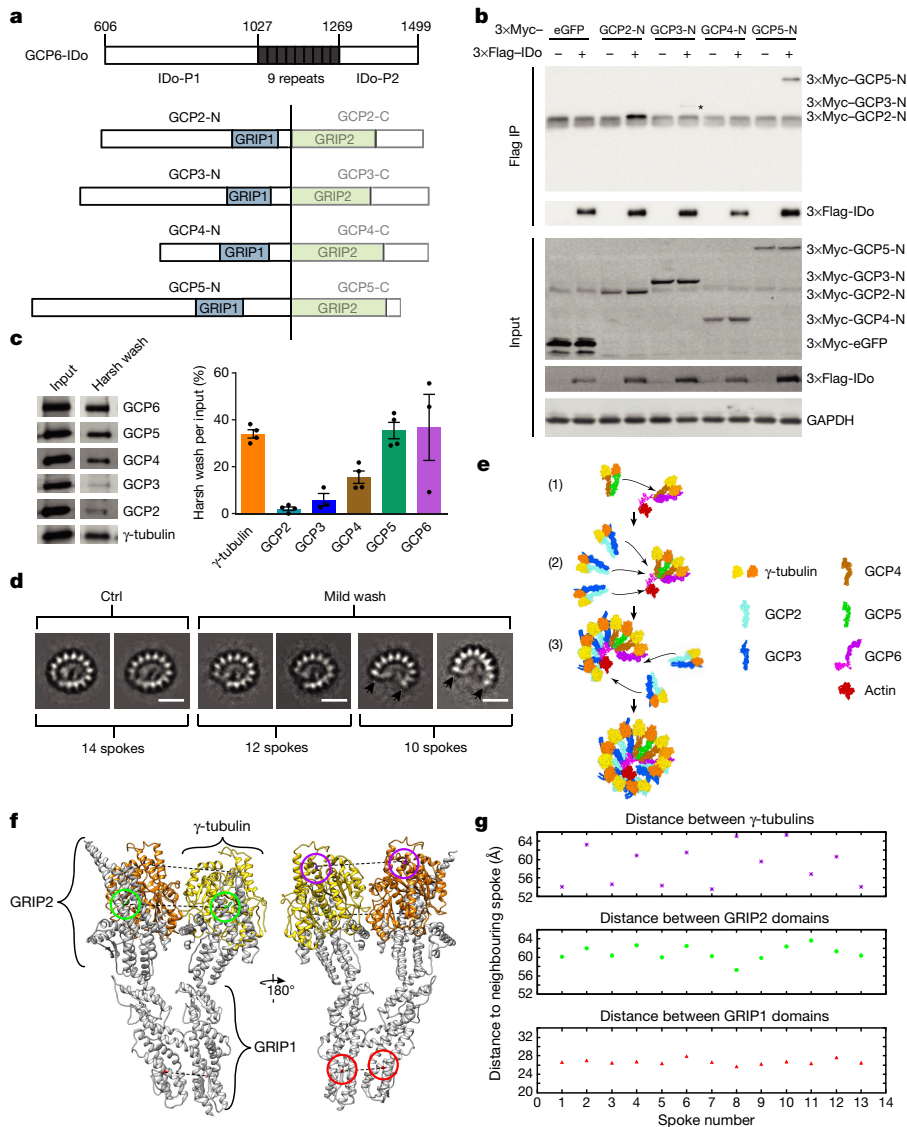


Fig. 3 | The GCP6 insertion domain assists in a stepwise assembly process of the γ -TuRC. **a, Protein constructs for immunoprecipitation. IDo, insertion domain. **b**, Co-immunoprecipitation (IP) of 3xMyc-GCP6 insertion domain with the 3xMyc-tagged N terminus of GCP proteins. 3xMyc-eGFP was used as a negative control. The indicated antibodies were used for immunoblotting. Result is representative of three independent experiments. Asterisk denotes faint band of 3xMyc-GCP3-N. For gel source data, see Supplementary Fig. 1. **c**, After high-salt treatment (harsh wash), fragmented γ -TuRCs were eluted and analysed by immunoblotting. Input was purified γ -TuRC. For gel source data, see Supplementary Fig. 1. Right, quantification of immunoblots; $n = 3$ (GCP3), 4 (γ -tubulin), 4 (GCP2), 4 (GCP4), 4 (GCP5) and 3 (GCP6) biologically independent experiments; data are mean \pm s.e.m. **d**, After mild salt treatment, fragmented γ -TuRCs were eluted and analysed by negative-stain electron microscopy. Representative 2D classes from three biologically independent experiments with similar results are shown. Black arrows point to the open ends. Scale bars, 20 nm. Ctrl, control. **e**, Potential stepwise assembly of the γ -TuRC. (1) Stable GCP4-GCP5 and GCP4-GCP6 heterodimers form the GCP4-GCP5-GCP4-GCP6 core. (2) Three GCP2-GCP3 heterodimers are recruited by interactions of N-GCP2 with the GCP6 insertion domain. (3) Two GCP2-GCP3 heterodimers bind to both ends of the 10-spoke intermediate. **f**, Three points were defined along the entire spoke based on the C α atoms of conserved amino acid residues (Ala in GRIP1 (red); Phe in GRIP2 (green); and Asn in γ -tubulin (magenta)). **g**, Distances between the points were computed for each spoke and its neighbour along the helical sequence. Colours are as in **f**.**

10 and 12) lacked distinguishing structural features of the GCP core segments, LFQ mass spectrometry (Fig. 2a) indicated that they correspond to the single copies of GCP5 and GCP6. Notably, we could not trace the GCP5 and GCP6 insertion domains (Extended Data Fig. 1a) in the cryo-EM reconstruction, which suggests that at least the segments directly associated with the GCP core fold are not well-ordered.

We next generated homology models for the core segments of *X. laevis* GCP2-GCP6 and *X. laevis* γ -tubulin using X-ray structures of human GCP4⁵ and γ -tubulin⁶, respectively and refined them against the cryo-EM density (Methods). The resulting atomic model had good statistics (Extended Data Table 1) and was validated against our cryo-EM density (Extended Data Fig. 2h). Consistent with the determined local resolution (Extended Data Fig. 2g), bulky amino acid side chains were resolved in many areas of the density and corresponded with the refined homology models (Extended Data Fig. 5, Methods). GCP-variant-specific bulky amino acid side chains confirmed our previous assignment of GCP identities for GCP2-GCP4 and also allowed identification of GCP5 (spoke 10) and GCP6 (spoke 12) in the cryo-EM density, thus establishing an atomic model for the structural core of the vertebrate γ -TuRC with a uniform sequence of GCPs: GCP(2-3)₄-GCP4-GCP5-GCP4-GCP6-(GCP2-3) (Fig. 1a). This order was unexpected because GCP4, GCP5 and GCP6 had previously been proposed to cap the γ -TuRC spiral^{3,7}.

Actin is an integral part of the γ -TuRC

Several density segments remained unexplained (Fig. 1a), and the most prominent was a belt of density lining the interior of the γ -TuRC that was already visible in the negatively stained γ -TuRC particles (Extended Data Fig. 1h, i). For assignment of these density segments, we pursued an unbiased structure-guided approach (Methods), which we verified by localizing γ -tubulin and GCP4 in the γ -TuRC as positive controls (Extended Data Fig. 6a) and ovalbumin as a negative control (Extended Data Fig. 6b). We were unable to localize established γ -TuRC components, including NEDD1 and NME7 that were identified by LFQ mass spectrometry (Fig. 2a), and MOZART1 (Extended Data Fig. 6c), which suggests that they are not covered in our cryo-EM reconstruction and probably associate with the unresolved GCP N termini. Many other proteins found to be abundant in the purified γ -TuRC could also not be localized (Methods). Unexpectedly, actin—also identified as a putative γ -TuRC component (Fig. 2a)—produced an unambiguous fit in a globular domain of the belt density (Fig. 1c, Extended Data Fig. 6d), where it bridges an α -helical belt domain with the γ -tubulin of spoke 2 (Fig. 1c). Actin binds to γ -tubulin via its D-loop—a canonical interaction also involved in the formation of actin filaments and DNaseI binding⁸. We confirmed the presence of actin by immunoblotting (Extended Data Fig. 6e) and showed that it colocalized with GCP6 by immunofluorescence microscopy (Extended Data Fig. 6f, g). Centrosomes have

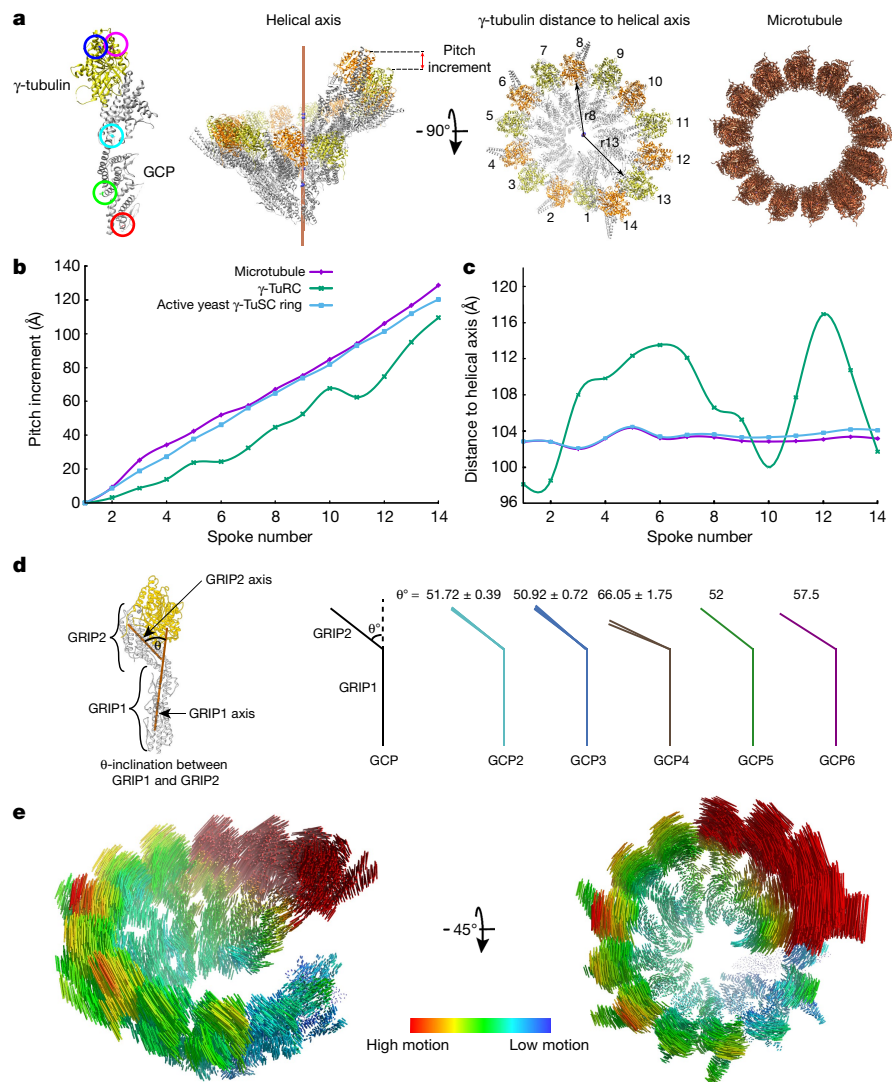


Fig. 4 | Geometrical and conformational analysis of the γ -TuRC. **a**, Residues distributed along each spoke (coloured circles) were used to define centroids of the approximate helical axis (brown). An MT cross-section is shown for comparison. **b**, Spokewise elevation along the helical axis (incremental pitch) for the γ -TuRC, the closed γ -TuSC spiral (PDB code 5FLZ) and a 13-spoked MT (PDB code 6EW0). **c**, Distance of γ -tubulins from the helical axis, plotted for the same complexes as in **b**. **d**, Inclinations between GRIP1 and GRIP2 axes, as

defined in the atomic model. Average inclinations for the different GCP variants; data are mean \pm s.e.m. $n = 5$ (GCP2), 5 (GCP3), 2 (GCP4), 1 (GCP5) and 1 (GCP6) according to the stoichiometry in the γ -TuRC. **e**, Conformational changes required for activation of the γ -TuRC visualized via vectors linking residues in both conformations. Vector length is colour-coded. See Supplementary Videos 1 and 2 for visualization of the conformational change.

previously been proposed to nucleate not only MTs but also actin filaments⁹ and we therefore tested whether γ -TuRC-associated actin could be involved. Although the ARP2/3-VCA complex used as control showed robust actin filament nucleation activity in vitro, the purified γ -TuRC was inactive (Extended Data Fig. 6h). We next used an inhibition experiment to test whether γ -TuRC-associated actin could have a role in MT nucleation. The actin-binding protein DNaseI binds the actin D-loop with high affinity, and thus competes with actin binding to the γ -tubulin of spoke 2⁸. DNaseI treatment of purified γ -TuRC significantly inhibited their MT nucleation activity in vitro, and pre-incubation of DNaseI with actin abolished this effect (Extended Data Fig. 6i, j). Thus, actin is a bona fide structural component of the γ -TuRC, and has functional relevance in MT nucleation.

GCP6 assists in assembly of the γ -TuRC

Candidates for representing the remaining mostly α -helical belt density are the insertion domains of GCP5 and GCP6. The GCP5 insertion

domain is comparably short (120 residues) and predicted to be mostly unordered (Extended Data Fig. 4c). By contrast, the 750-residue-long GCP6 insertion domain contains a segment of 249 residues predicted to be highly α -helical (approximately 70%) (Extended Data Fig. 4d), which we confirmed by circular dichroism spectroscopy (Extended Data Fig. 4e). Furthermore, the length of this GCP6 segment is in good agreement with the size of the α -helical belt domain (Methods), rendering the GCP6 insertion domain a genuine candidate for representing this part of the cryo-EM density.

Our cryo-EM density suggests a direct interaction between the α -helical belt domain and the N-terminal region of various GCP proteins. Following this observation, we analysed whether the human GCP6 insertion domain (residues 606–1499) (Fig. 3a) has the ability to interact with N-terminal domains of human GCP proteins. Indeed, the N-terminal domains of GCP2 and GCP5 (GCP2-N and GCP5-N, respectively) were robustly co-immunoprecipitated with the GCP6 insertion domain (Fig. 3b), whereas GCP3-N showed only weak binding (Fig. 3b, asterisk). Enhanced green fluorescent protein (eGFP) that was

used as control and GCP4-N were not detected. To narrow down which regions of the GCP6 insertion domain interact with GCP-N termini, we divided the human GCP6 insertion domain into three subdomains (Fig. 3a). GCP2-N and GCP5-N interacted specifically with the mostly conserved α -helical GCP6(606–1026) fragment, denoted part 1 of the GCP6 insertion domain, which consequently could mediate specific recruitment of GCP2 and GCP5 to the γ -TuRC (Extended Data Figs. 7a, 8). Mutations that disrupt the α -helical structure (Extended Data Fig. 8) strongly reduced binding to GCP2-N and GCP5-N (Extended Data Fig. 7b), which validates the role of the GCP6 insertion domain for GCP recruitment. In contrast to the α -helical region, the nine repeats in the GCP6 insertion domain did not interact with GCP2 and GCP5 (Extended Data Fig. 7a). Differential extraction of GCP variants from the purified γ -TuRC by salt treatment suggests a stable core of GCP4, GCP5 and GCP6 that is resistant to harsh salt treatment (Fig. 3c). Mild salt treatment depleted γ -TuRC of peripheral GCP subunits (most likely GCP2–GCP3 complexes), as indicated by negative-stain 2D classes (Fig. 3d). These data suggest a stepwise assembly mechanism in which the GCP6 insertion domain mediates specific recruitment of one pre-assembled GCP4–GCP5 dimer¹⁰ to a GCP4–GCP6 core¹¹, before the binding of preformed GCP2–GCP3 complexes (Fig. 3e). The presence of preformed GCP2–GCP3 complexes is not only supported by their concomitant loss after a salt wash always resulting in even-spoked γ -TuRCs (Fig. 3d), but also by the observation of a pairwise pattern in the distances of GCP and γ -tubulin molecules, indicating tighter interaction within the GCP2–GCP3 dimers compared with interdimer interactions (Fig. 3f, g).

Conformational activation of the γ -TuRC

In vertebrates, the γ -TuRC is assembled in a state with only basal MT nucleation activity and activators are required to stimulate MT nucleation^{12,13}. To gain insights into the structural basis of γ -TuRC activation, we explored whether the geometry of the γ -TuRC in our cryo-EM reconstruction would be compatible with templating MT nucleation. We approximated the γ -TuRC helical axis (Fig. 4a) and, for each spoke, determined the incremental γ -tubulin elevation along the axis (Fig. 4b) and the helix radius (Fig. 4c). We observed no strict helical symmetry for both parameters, which indicates that the purified γ -TuRC was geometrically incompatible with being a structural template for a 13-spoked MT (Fig. 4b, c). To understand the structural basis for the observed deviations in helical symmetry, we analysed the intrinsic conformational arrangement of GRIP1 and GRIP2 domains for each individual spoke. The domain arrangement was similar for all copies of GCP2, GCP3 and GCP5, but differed for GCP4 and GCP6 (Fig. 4d). Here, the GRIP domains were arranged in a more-stretched conformation, resulting in position-specific displacement of the γ -tubulins that could contribute to the observed distortion of helical symmetry. We extrapolated the structure of an MT-nucleation-competent γ -TuRC based on the previously described *Saccharomyces cerevisiae* γ -TuSC spiral¹⁴ known to have high MT nucleation activity, and visualized the required global conformational rearrangement, mostly representing a contraction of the γ -TuRC spiral that is accompanied by repositioning of γ -tubulins to achieve a uniform spacing that reflects MT symmetry (Fig. 4e, Supplementary Videos 1, 2). Such structural rearrangements could occur spontaneously during MT nucleation, explaining the basal nucleation activity of purified γ -TuRC in vitro (Extended Data Fig. 1c), or they could be induced by activators of MT nucleation such as CEP215¹². To test the latter hypothesis, we added the purified CEP215 N terminus (CEP215-N) including the activating CM1 motif in large excess to purified γ -TuRC, and analysed the structural and functional properties of the complexes using negative-stain electron microscopy and in vitro

MT nucleation assays (Extended Data Fig. 9a, b). Although CEP215-N bound to γ -TuRC in vitro (Extended Data Fig. 9c), we did not observe structural changes or a strong increase in MT nucleation activity, which suggests the requirement of other factors that potentially act together with CEP215-N. Consistently, the addition of recombinant CEP215-N to egg extract stimulated MT nucleation only in combination with Ran(Q69L) (loaded with GTP)—a GTP-hydrolysis-defective mutant of the GTPase Ran (Extended Data Fig. 9d). Notably, the N terminus of the CEP215(F75A) mutant¹² that was less efficient in γ -TuRC binding (Extended Data Fig. 9c) could not stimulate MT nucleation activity in egg extract (Extended Data Fig. 9d).

Collectively, these biochemical and structural data fundamentally deepen our understanding of MT nucleation by providing detailed insights into the molecular organization, assembly and activation mechanism of vertebrate γ -TuRC. Our study provides a rationale for the evolutionary acquisition of the GCP4–GCP6 variants in vertebrates. In contrast to budding yeast, in which the MT nucleating template is assembled only at the yeast centrosome for local MT nucleation¹⁴, our data suggest that the acquisition of GCP4 and GCP6—both contributing to the asymmetry of γ -TuRC—enables context-independent pre-assembly of inactive γ -TuRC that can be efficiently activated via a conformational change. This allows much faster and more dynamic regulation of MT nucleation within the assembling mitotic spindle, in which hundreds to thousands of MTs have to be nucleated within minutes¹⁵.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1896-6>.

- Borisy, G. et al. Microtubules: 50 years on from the discovery of tubulin. *Nat. Rev. Mol. Cell Biol.* **17**, 322–328 (2016).
- Zheng, Y., Wong, M. L., Alberts, B. & Mitchison, T. Nucleation of microtubule assembly by a γ -tubulin-containing ring complex. *Nature* **378**, 578–583 (1995).
- Kollman, J. M., Merdes, A., Mourey, L. & Agard, D. A. Microtubule nucleation by γ -tubulin complexes. *Nat. Rev. Mol. Cell Biol.* **12**, 709–721 (2011).
- Schmidt-Cernohorska, M. et al. Flagellar microtubule doublet assembly *in vitro* reveals a regulatory role of tubulin C-terminal tails. *Science* **363**, 285–288 (2019).
- Guillet, V. et al. Crystal structure of γ -tubulin complex protein GCP4 provides insight into microtubule nucleation. *Nat. Struct. Mol. Biol.* **18**, 915–919 (2011).
- Aldaz, H., Rice, L. M., Stearns, T. & Agard, D. A. Insights into microtubule nucleation from the crystal structure of human γ -tubulin. *Nature* **435**, 523–527 (2005).
- Moritz, M., Braunfeld, M. B., Guénebaud, V., Heuser, J. & Agard, D. A. Structure of the γ -tubulin ring complex: a template for microtubule nucleation. *Nat. Cell Biol.* **2**, 365–370 (2000).
- Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F. & Holmes, K. C. Atomic structure of the actin:DNase I complex. *Nature* **347**, 37–44 (1990).
- Farina, F. et al. The centrosome is an actin-organizing centre. *Nat. Cell Biol.* **18**, 65–75 (2016).
- Farache, D. et al. Functional analysis of γ -tubulin complex proteins indicates specific lateral association via their N-terminal domains. *J. Biol. Chem.* **291**, 23112–23125 (2016).
- Anders, A., Lourenço, P. C. & Sawin, K. E. Noncore components of the fission yeast γ -tubulin complex. *Mol. Biol. Cell* **17**, 5075–5093 (2006).
- Choi, Y. K., Liu, P., Sze, S. K., Dai, C. & Qi, R. Z. CDK5RAP2 stimulates microtubule nucleation by the γ -tubulin ring complex. *J. Cell Biol.* **191**, 1089–1095 (2010).
- Scrofani, J., Sardon, T., Meunier, S. & Vernos, I. Microtubule nucleation in mitosis by a RanGTP-dependent protein complex. *Curr. Biol.* **25**, 131–140 (2015).
- Kollman, J. M., Polka, J. K., Zelter, A., Davis, T. N. & Agard, D. A. Microtubule nucleating γ -TuSC assembles structures with 13-fold microtubule-like symmetry. *Nature* **466**, 879–882 (2010).
- Petry, S. & Vale, R. D. Microtubule nucleation at the centrosome and beyond. *Nat. Cell Biol.* **17**, 1089–1093 (2015).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

Article

Methods

Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

Plasmid construction

PCR amplifications were performed with Q5 High Fidelity DNA Polymerase (NEB). Fragments were inserted into the backbones with NEBuilder HiFi DNA Assembly Cloning Kit (NEB). To generate 3×Flag-tagged GCP6 constructs, corresponding fragments for the GCP6 insertion domain (residues 606–1499, denoted GCP6-IDo), the 4P mutant of the GCP6 insertion domain (denoted GCP6-IDo^{4P}), the GCP6 insertion domain part 1 (residues 606–1026, denoted IDo-P1), the 4P mutant of the GCP6 insertion domain part 1 (denoted IDo-P1^{4P}), the GCP6 insertion domain 9 repeats (residues 1027–1268, denoted IDo-9 repeats) and the GCP6 insertion domain part 2 (residues 1269–1499, denoted IDo-P2) were amplified and inserted into BamHI-digested pRetroX-TRE3G vector and sub-cloned into pCMV-3Tag-1A vector via BamHI. To generate 3×Myc-tagged GCP6 constructs, eGFP, GCP2-N (1–1485 bp), GCP3-N (1–1620 bp), GCP4-N (1–1008 bp) and GCP5-N (1–2142 bp) were amplified and inserted into the BamHI site of pCMV-3Tag-2B. An additional base pair 'A' was introduced upstream to avoid frameshift and the BamHI site was regenerated downstream. The coding sequence for residues 546–794 of *X. laevis* GCP6 (xGCP6(546–794)) was amplified and inserted into a BamHI-digested pGEX-6P-1 vector to generate pGEX-6P1-xGCP6(546–794). Subsequently, the plasmid was digested with EcoRI/NotI and a StrepII tag was introduced after the coding sequence of xGCP6(546–794).

Antibodies

Anti-γ-tubulin rabbit polyclonal antibody, which was used for γ-TuRC purification and the CEP215-N pull-down assay, was generated against the C-terminal peptide². Anti-γ-tubulin mouse monoclonal antibody (GTU-88), which was used for immunoblotting, was from Sigma-Aldrich. Rabbit anti-GCP2 polyclonal antibody was from Thermo Fisher Scientific. Rabbit anti-GCP3 polyclonal antibody and rabbit anti-GCP6 polyclonal antibodies were from Y. Zheng. Guinea-pig anti-GCP6 polyclonal antibody for immunofluorescence was generated as previously described¹⁶. Anti-β-actin mouse monoclonal antibody (AC-74) used in immunofluorescence and anti-actin rabbit polyclonal antibody (A2066) used in immunoblotting were from Sigma-Aldrich. Anti-GCP4 rabbit polyclonal antibodies were raised against full-length purified GCP4. Anti-GCP5 mouse monoclonal antibody (E-1) was from Santa Cruz Biotechnology. Mouse anti-Flag monoclonal antibody (9A3) and rabbit anti-GAPDH polyclonal antibody (14C10) were from Cell Signaling Technology. Mouse monoclonal anti-Myc antibody (clone 9E10) was from Sigma-Aldrich. Secondary antibodies were: donkey anti-mouse Alexa Fluor 488-conjugated antibody and goat anti-guinea pig Alexa Fluor 555-conjugated antibody (Thermo Fisher Scientific); peroxidase-conjugated goat anti-mouse antibody (Jackson ImmunoResearch Laboratories); donkey anti-mouse DyLight 680 and 800-conjugated antibodies (Thermo Fisher Scientific); anti-rabbit DyLight 680-conjugated antibody (Cell Signaling Technology); IRDye 800CW Donkey anti-Rabbit IgG (LI-COR Biosciences).

γ-TuRC purification

CSF-arrested *X. laevis* egg extracts were prepared as previously described¹⁷ and stored at –80 °C. To purify the γ-TuRCs, an adequate amount of egg extracts was defrosted and incubated with γ-tubulin antibody-crosslinked Dynabeads Protein A for 30 min at room temperature and subsequently washed with: (1) 3 times with CSF-XB buffer (5 mM EGTA, 10 mM HEPES pH7.7, 2 mM MgCl₂, 50 mM sucrose, 100 mM KCl, 0.1 mM CaCl₂); (2) 3 times with CSF-XB buffer supplemented with

250 mM KCl and 0.3% Triton X-100; and (3) twice with HB100 buffer (50 mM Na-HEPES pH 8.0, 1 mM EGTA, 1 mM MgCl₂, 100 mM NaCl) containing 0.1 mM GTP. Elution was performed overnight at 4 °C with gentle rotation in HB100 buffer containing 0.1 mg ml^{–1} γ-tubulin antigenic C-terminal peptide², 1 mM GTP and 0.02% Tween20. The concentration of the purified γ-TuRCs was determined as 5 nM by immunoblotting comparing to purified human γ-tubulin.

γ-TuRC salt treatments

Both mild and harsh washing were performed with the same protocol as γ-TuRC purification except washing step (2). For the mild washing, washing step (2) was carried out with CSF-XB buffer supplemented with 500 mM KCl and 0.3% Triton X-100 (3 times). For harsh washing, washing step (2) was carried out sequentially with CSF-XB buffer containing 0.3% Triton X-100 and 4 different concentrations of additional KCl: 250 mM (4×), 500 mM (1×), 750 mM (1×) and 1 M (1×). The beads were washed with each buffer for 15–20 min (at 4 °C with gentle rotation) and the proteins were eluted as described in the γ-TuRC purification protocol. The band intensity was quantified with the software Image Studio Lite (v.5.2.5, LI-COR Biosciences) using the Analysis function.

In-gel tryptic digestion, LC-MS/MS analysis and database search

Samples were separated by SDS-PAGE (minigel, 10%; Invitrogen) for 1.0 cm. Coomassie-stained lanes were cut out with a scalpel and processed as previously described¹⁸. In brief, samples were reduced with dithiothreitol (DTT), alkylated with iodoacetamide and digested with trypsin. Peptides were extracted from gel pieces, concentrated in a speedVac vacuum centrifuge and dissolved with 15 μl 0.1% trifluoroacetic acid. Nanoflow LC-MS2 analysis was performed with an Ultimate 3000 liquid chromatography system coupled to an Orbitrap Elite mass spectrometer (Thermo-Fisher). Five microlitres of sample was injected to a self-packed analytical column (75 mm × 200 mm; ReproSil Pur 120 C18-AQ; Dr. Maisch) and eluted with a flow rate of 300 nl min^{–1} in an acetonitrile-gradient (3–40%). One survey scan (resolution: 60,000) was followed by 15 information-dependent product ion scans in the ion trap. The MaxQuant software (1.6.2.6a)¹⁹ was used with default settings for database search against a *X. laevis* database downloaded from UniProt.org (Proteome ID: UP000186698 with 42,878 entries; last modified October 2018) together with the contaminants database included in the MaxQuant software. In addition, a custom-made database (Supplementary Table 2) was used containing additional database entries for γ-tubulin complex components not identical with the UniProt entries. Trypsin was specified as enzyme. Carbamidomethyl was set as fixed modification of cysteine and oxidation (methionine), deamidation (asparagines and glutamine) and N-terminal acetylation as variable modifications. A false-discovery rate of 1% was used on peptide and protein levels. To calculate iBAQ values²⁰, an additional MaxQuant analysis was used with iBAQ calculation enabled. An Andromeda score threshold of 40 was set for unmodified peptides to avoid false positives for quantification. For calculation of stoichiometries, the calculated iBAQ values of the database entries Q5PQ98 (γ-tubulin), O73787 (GCP3), A0A1L8HOR3 (GCP4), A0A1L8HGZ5 (GCP5), A0A1L8GZ56 (GCP6), A0A1L8GY92 (NEDD1), A0A1L8EXC8 (actin) and A0A1L8HF79 (NME7) from UniProtKB and XP_018080016.1 (GCP2) from the NCBI database were used. To account for different degrees of purity among different preparations of γ-TuRC, data were normalized to the iBAQ values of GCP2 and GCP3. Source data for LFQ mass spectrometry are included in Supplementary Table 1.

Immunoprecipitation

Mycoplasma-free HEK293T cells were bought from ATCC (American Type Culture Collection) and grown in Gibco DMEM/F-12 medium (Thermo Fisher Scientific) supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 100 U ml^{–1} penicillin and 100 μg ml^{–1} streptomycin at 37 °C with 5% CO₂. Plasmids were transiently transfected at 40–50%

cell confluency in 10-cm dishes with polyethylenimine according to the standard protocol. Then, 18–36 h after transfection, cells were scraped and lysed with lysis buffer containing 10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.5% NP40, 0.5 mM EDTA, 1 tablet per 50 ml protease inhibitor cocktail (cOmplete, EDTA-free, Roche) and 1:500 Benzamide (Merck Millipore). After pipetting 20 times and passing through a 21Gx1½ (0.8 × 40 mm) needle 20 times, the mixture was placed on ice for 10 min and clarified by spinning for 20 min, 20,000g at 4 °C. The supernatant was collected and incubated with 25–30 µl anti-Flag M2 affinity gel (Sigma-Aldrich, equilibrated with the lysis buffer) for 2 h at 4 °C on a rotating wheel. The beads were collected by centrifuging at 6,000g, 4 °C for 30 s. After washing once with lysis buffer and three times with PBS, the proteins on the beads were eluted by incubating with 2× sample buffer at 65 °C for 15 min.

Silver staining

Polyacrylamide 10–12% gels were used in this study. For silver staining, the gel was incubated for 30 min in fixing solution containing 40% methanol and 10% acetic acid. Excess acetic acid was washed away with 30% methanol followed by three washes with water. The gel was sensitized for 2 min in 0.02% freshly prepared Na₂S₂O₃ and then washed with water. The staining was performed with 0.2% AgNO₃ at 4 °C for 25 min, and residual AgNO₃ was washed away with water. The development was performed with 6% Na₂CO₃, 0.05% formaldehyde and 0.0004% Na₂S₂O₃, and the reaction was terminated with 1.4% EDTA when the staining was sufficient. Major bands from the gel in Extended Data Fig. 1d were cut out and verified by mass spectrometry.

In vitro MT nucleation assay

Unlabelled and Cy3-labelled pig brain tubulin were mixed 24:1 in 1× BRB80 buffer (80 mM PIPES/KOH pH 6.8, 1 mM MgCl₂ and 1 mM EGTA) with a final glycerol concentration of 12.5% (w/v). After spinning at 352,860g, 4 °C for 5 min with S100-AT3 rotor (Thermo Fisher Scientific), 2.5 µl supernatant and 2.5 µl γ-TuRC mixture (0.2 µl γ-TuRC, 1 mM GTP, 12.5% (w/v) glycerol in 1× BRB80 buffer) were used, incubated for 15–30 min on ice and transferred to a 37-°C water bath for 5 min for MT nucleation. The sample was immediately mixed with 50 µl 1% glutaraldehyde in 1× BRB80 buffer and incubated for 5 min at room temperature for crosslinking. Then, 1 ml cold 1× BRB80 buffer was added to stop the reaction and 50 µl of sample was mounted onto a 2 ml 10% glycerol, BRB80 cushion in a Corex 15-ml glass tube with a 12-mm coverslip supported by a glass platform at the bottom. The MTs were sedimented onto poly-lysine-coated coverslips by centrifuging with an HB4/HB6 rotor (Thermo Fisher Scientific) at 23,530g for 1 h at 20 °C. After fixing for 5 min in ice-cold methanol, the sample was mounted on a 2 µl drop of Citifluor AF1 (Electron Microscopy Sciences) and sealed with nail polish. For each sample, 20 random images were acquired using an Axiovert 200 M microscope (Carl Zeiss Microscopy) equipped with a Plan-Apochromat 63× NA 1.3 oil objective lens (Carl Zeiss Microscopy), and a Cascade1K EMCCD camera (Photometrics). Imaging was operated with VisiView and the images were processed with ImageJ.

DNaseI inhibition assay

Purified γ-TuRC was incubated with a 100-fold excess of recombinant DNaseI (Sigma-Aldrich, D5319) for 3 h on ice. As a control, DNaseI and G-actin were pre-incubated in a 1:1 ratio for 1 h on ice, and the mixture was subsequently added to the purified γ-TuRC. The protocol for the in vitro MT nucleation assay was performed as described above except: after the addition of 1 ml cold 1× BRB80 buffer to stop the reaction, 3 µl of sample was squash-fixed on a slide with a 12 × 12-mm² coverslip. The imaging protocol was as described above and the MT number was counted manually from 20 random fields. For each sample, five random images were compiled into a stack and the maximum projection was presented in Extended Data Fig. 9b. This protocol is the same for the in vitro MT nucleation assay of γ-TuRC with CEP215-N.

Sucrose gradient

Sucrose gradients were made by a Model 106 Gradient Master (BioComp Instruments) according to standard instructions. Purified γ-TuRC was loaded onto a 2.2-ml 5–40% gradient in HB100 buffer with 0.1 mM GTP. Centrifugation was performed in PA 7/16 X 2-3/8 tubes (Beckman Coulter) in the S55-S Swinging-Bucket Rotor (Thermo Fisher Scientific) at 200,000g for 3 h at 4 °C. The fractions were collected from the top (160 µl per fraction). In total, 13 fractions were collected and analysed by immunoblotting. γ-TuRC was found around fractions 8–11. Thyroglobulin 19S (660 kDa) was used as a molecular mass marker.

Protein purification

For expression of the N-terminal 249 amino acids of the xGCP6-IDo, pGEX-6P1-xGCP6(546–794)-Strep II was transformed into *Escherichia coli* strain BL21-CodonPlus(DE3)-RIL competent cells. The transformed cells were cultured in 2xYT medium, and protein expression was induced with 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at an optical density of 0.5–0.8. After induction for 5 h at 25 °C, cells were collected, washed with PBS and stored at –80 °C. The pellets were resuspended in lysis buffer containing 50 mM Tris, 150 mM NaCl, 10% glycerol, pH 8, 1 mM EDTA, 1 mM DTT, 5 mM ATP, 2 mM MgCl₂, 0.1% Triton X-100, cOmplete EDTA-free protease inhibitor cocktail (Roche) and PMSF. Afterwards, the cells were lysed by sonication. The cell lysate was clarified by centrifuging at 45,000 rpm for 25 min with Type 50.2 Ti Rotor (Beckman Coulter). The supernatant was incubated with the pre-equilibrated Protino Glutathione Agarose 4B (MACHEY-NAGEL), and the protein was eluted by 3C protease cleavage. The eluates were bound to StrepTactin TM Sepharose High Performance (GE Healthcare) and eluted with 2.5 mM desthiobiotin.

For purification of GST, GST–CEP215-N and GST–CEP215(F75A)-N, the protocol is as described for xGCP6(546–794) except the induction and elution steps. Protein expression was induced overnight at 18 °C, and after the purification washing steps, proteins were eluted with 30 mM reduced L-glutathione (Sigma-Aldrich) in 50 mM Tris, 150 mM NaCl, 10% glycerol, pH 8 and 1 mM EDTA. Eluted proteins were further purified using a Mono Q 5/50 GL anion exchange column (GE Healthcare) to remove the reduced L-glutathione.

Circular dichroism spectroscopy

Purified xGCP6(546–794) was dialysed against 50 mM Na₂HPO₄·NaH₂PO₄ and 150 mM NaF, pH 8 overnight and concentrated on Vivaspin 6 centrifugal concentrator (Sartorius) to a final concentration of 2.6 µM. The circular dichroism spectra were recorded on a Jasco J715 spectropolarimeter at 25 °C. A cell with a 1-mm path length was used for spectra recorded between 190 and 250 nm with sampling points every 0.2 nm. For each measurement, the spectra represented the average of five scans and circular dichroism intensities were expressed in millidegrees.

CEP215-N pull-down assay

Purified GST, GST–CEP215-N and GST–CEP215(F75A)-N were conjugated with Protino Glutathione Agarose 4B (MACHEY-NAGEL) and pre-equilibrated with HB100 buffer with 0.1 mM GTP. Seven and a half microlitres purified γ-TuRC was incubated with the conjugated beads in a volume of 90 µl for 2 h at 4 °C. After washing 4 times with washing buffer (HB100 buffer with 0.1 mM GTP and 0.01% NP-40), bound proteins were eluted with elution buffer (HB100 buffer with 0.1 mM GTP, 0.01% NP-40 and 30 mM reduced GSH) at 4 °C for 30 min.

CEP215-N activity assay in egg extract

Freshly prepared CSF-arrested *X. laevis* egg extracts were mixed with 1 µM Cy3-labelled pig brain tubulin, 15 µM Ran(Q69L), 3 µM GST–CEP215-N or GST–CEP215(F75A)-N and incubated on ice for 1 h. After incubating in a 20-°C water bath for 15 min, 2 µl of sample was squashed on a slide with a 12 × 12-mm² coverslip to test for aster formation. Ten random images

Article

were acquired for each sample using an Axiovert 200 M microscope (Carl Zeiss Microscopy) equipped with a 10× objective lens (Carl Zeiss Microscopy), and a Cascade1 K EMCCD camera (Photometrics). Imaging was operated with VisiView. The images were processed with ImageJ and the fluorescence intensity of the asters in 10 images was quantified.

Immunofluorescence microscopy

Purified γ -TuRCs from *X. laevis* were spun on 12-mm coverslips coated with poly-L-lysine (Sigma-Aldrich) at 23.530g, 4 °C for 30 min with an HB6 rotor (Thermo Fisher Scientific). After fixation with ice-cold methanol for 5 min at −20 °C, samples were incubated with PBS containing 10% FBS and 0.1% Triton X-100 for 30 min at room temperature and subsequently treated with 1% SDS in PBS for 5 min at room temperature. The PBS-washed coverslips were sequentially incubated at room temperature with primary antibodies for 1 h and Alexa-Fluor-conjugated secondary antibodies for 30 min. A DeltaVision RT system (Applied Precision, Olympus IX71 based) equipped with the Photometrics CoolSnap HQ camera (Roper Scientific), a 100×/1.4 NA UPlanSAPO objective (Olympus), a mercury arc light source and the softWoRx software (Applied Precision) were used for imaging. Selected channels, FITC and TRITC, with different exposure times according to the fluorescence intensity of each protein, were applied. Images were acquired by softWoRx software (Applied Precision) and analysed by Fiji. Each group included at least 15 images in each independent experiment. For quantification, one randomly selected image was used as a reference for adjusting the brightness and contrast levels to optimal values and all images were analysed using the same settings. GCP6 and actin signals from images in each experiment were counted and used for quantification.

Actin polymerization assay

Pyrene-labelled actin (Hypermol) was dissolved in general actin buffer (5 mM Tris, pH 8.0, 0.2 mM CaCl_2 , 0.2 mM ATP and 0.5 mM DTT) to a final concentration of 2 μM . The reaction contained 20 μl pyrene-labelled actin, actin polymerization buffer (final concentration with 10 mM Tris, pH 7.5, 50 mM KCl, 2 mM MgCl_2 and 1 mM ATP) and either 2 μl purified γ -TuRC (final concentration 0.5 nM) or 2 μl γ -TuRC elution buffer. The controls contained pyrene-labelled actin, actin polymerization buffer, γ -TuRC elution buffer, Arp2/3 (0.5 nM final concentration, Hypermol) and VCA (15 nM, Hypermol). Fluorescence was measured at 25 °C using a plate reader every 1 min for a period of 100 min (CLARIOstar, BMG Labtech, excitation, F:360-10, emission, F: 450-10).

Negative-staining electron microscopy and image analysis

Negatively stained samples were generated as follows: 5 μl of the purified γ -TuRCs were applied on a glow-discharged copper-palladium hexagonal 400 EM mesh grid covered with an approximately 10-nm-thick continuous carbon layer. After a 30 s incubation, the sample was blotted on a Whatman filter paper 50 (1450-070) and washed with three drops of water. Samples on grids were stained with 3% uranyl acetate in water. Images were acquired on an FEI Tecnai F20 electron microscope, operated at 200 kV, equipped with a field emission gun and bottom-mounted 4 K camera. The micrographs were acquired at 50,000× magnification by the SerialEM software, resulting in 2.27 Å per pixel. For 2D classification, γ -TuRC particles were selected manually using the Boxer of EMAN2²¹. Image processing was carried out using the IMAGIC-4D package²². Particles were band-pass-filtered and normalized in their grey value distribution, and mass-centred. Two-dimensional alignment, classification and iterative refinement of class averages were performed as previously described²³. Approximate number of total particles for 2D classification and averaging: Fig. 3d, ctrl, 5,400; Fig. 3d, mild wash, 3,100; both Extended Data Fig. 1h, i, 2,200. For Fig. 3d and Extended Data Fig. 1h, a 46.5-nm mask was used for averaging; for Extended Data Fig. 1i, a 14.5-nm mask was used for averaging. Particles included in the representative classes of Fig. 3d: 14 spokes, left, 115, right, 139; 12 spokes, left, 67, right, 59; 10 spokes, left, 76, right, 69.

For γ -TuRCs incubated with or without CEP215-N, the negative-staining sample preparation and imaging were done as described above. For 3D classification and 3D reconstruction of γ -TuRC with or without CEP215-N, particles were manually located on the micrographs as described above. Micrographs and particle coordinates were subsequently imported into Relion 3.0 beta and particles were subjected to 3D classification into three classes using a scaled density of the γ -TuRC cryo-EM reconstruction as initial reference. For the γ -TuRC without CEP215-N, we retained all 2,205 manually selected particles and subjected them to 3D autorefinement. For the γ -TuRC with CEP215-N, we retained 1,581 of the 2,057 manually selected particles for 3D autorefinement. To compare the overall structure and conformation of the two resulting 3D reconstructions, we superposed them according to the first four spokes using the Fit in map command in UCSF Chimera. To compare the structure of the CEP215-N- γ -TuRC density with the extrapolated structure of the active γ -TuRC, we simulated the density at 30 Å resolution from the atomic model in UCSF Chimera and superposed it to the negative stain 3D reconstruction of the CEP215-N- γ -TuRC complex as described above.

Cryo-EM sample preparation

Homemade graphene oxide holey carbon grids (Cu R2/1; 300 mesh) were glow-discharged using a Gatan Solarus 950 plasma cleaner for 20 s. Cryo-EM grids were prepared using a Thermo Fisher/FEI Vitrobot Mark IV operated at 22–25 °C and 60–70% humidity. Four microlitres of purified γ -TuRCs was applied to the EM grids within the climate chamber of the Vitrobot. After a waiting time of 30 s, grids were blotted with Whatman filter paper no. 1 for 5 or 10 s and plunge-frozen in liquid ethane bath cooled by liquid nitrogen.

Cryo-EM data acquisition

Four datasets were acquired using SerialEM²⁴ on a Titan Krios TEM (Thermo Fisher/FEI) operated at 300 kV and equipped with a K3 camera (Gatan) operated in dose fractionation mode. Datasets 1–3 were collected at an object pixel size of 2.1 Å per pixel with cumulative doses of 35 $e^-/\text{Å}^2$ (70 frames), 42 $e^-/\text{Å}^2$ (60 frames) and 51 $e^-/\text{Å}^2$ (46 frames), respectively. Dataset 4 was collected at an object pixel size of 1.35 Å per pixel with a cumulative dose of 57 $e^-/\text{Å}^2$ (39 frames). For each preselected hole, defocus was adjusted automatically before acquisition of two (datasets 1–3) or four (dataset 4) frame stacks per hole. Data were collected in a defocus range of −0.5 to −3.5 μm .

Data processing

All image processing steps are summarized in Extended Data Fig. 2. Unless stated otherwise, image processing was carried out in RELION 3.0-beta²⁵. Datasets 1–3 were initially processed separately, but following an identical workflow, which is described below. The individual numbers of frame stacks and particles retained after each processing step, as well as resolution of intermediary cryo-EM densities is included in Extended Data Fig. 2. Frame stacks were motion-corrected with Motion-Corr²⁶ using 5 × 5 patches. The contrast transfer function (CTF) of the motion-corrected micrographs was estimated using gCTF²⁷. In total, 3,000 particles were manually selected and used to generate a purely data-driven initial 3D density of the γ -TuRC in Relion with standard parameters. Particles were autopicked in RELION using this initial 3D density, from which reference projections were automatically generated. Localized particles were extracted at 4.2 Å pixel size in boxes of 128 × 128 pixels. Particles were subjected to two rounds of 3D classification with standard parameters to successively remove false-positive particles, broken γ -TuRCs and classes with strong orientational bias. The subset of particles retained after the second round of 3D classification were recentred, extracted at full spatial resolution in boxes of 256 × 256 pixels and subjected to 3D autorefinement using solvent-flattened Fourier shell correlation (FSC) and otherwise standard parameters. After refinement, particles were subjected to per-particle CTF refinement

(including beam tilt estimation for individual datasets) and Bayesian polishing²⁸ trained on 1,000 particles.

Dataset 4 was the largest, and autopicked particles were divided into four different subsets to speed up initial 3D classification steps. However, following the classification strategy described above, 3D classes of sufficient quality could only be obtained for two of the four subsets as the ratio of true- to false-positive particles after autopicking was much lower than for datasets 1–3. This was owing to the generally lower quality of the γ -TuRC purification for this dataset and the smaller field of view—both resulting in a very low number of particles on each micrograph. To retrieve particles from the two subsets of data that could not be classified successfully using the standard approach, we added particles belonging to the high-quality 3D classes from the other two subsets as ‘nucleators’ for 3D classes with sufficient quality but removed them afterwards. Merged particles from all four subsets were subjected to two additional rounds of 3D classification to retrieve the final set of particles from dataset 4. These particles were extracted at 2.1 Å pixel size in boxes of 256 × 256 pixels and then processed identically to datasets 1–3.

A total of 46,096 polished particles from all four datasets were merged and subjected to 3D autorefinement, which then served as a basis for several rounds of 3D multibody refinement²⁹. Using the cryo-EM density from 3D autorefinement, we prepared two sets of shape masks either (i) dividing the γ -TuRCs into seven dimers of successive spokes, or (ii) splitting the density into four segments, in which two segments represent the spoke ‘heads’ (C-terminal part of GRIP2 plus γ -tubulin) for spokes 1–7 and 8–14, respectively, and two segments represent the spoke ‘base’ (GRIP1 and N-terminal part of GRIP2) for spokes 1–7 and 8–14, respectively. All masks were generated by choosing an appropriate density threshold level for each density segment and extending the density envelope by 5 pixels and a soft edge of 5 pixels. Both sets of masks were used for a separate round of 3D multibody refinement. While 3D multibody refinement using the first set of masks resulted in much improved density for the spoke heads, the second set of masks yielded optimal refinement for the spoke bases. The resulting unfiltered density segments were merged into a composite map using UCSF Chimera³⁰. Global resolution of the final density was estimated 4.9 Å; however, local resolution estimation indicated significantly lower resolution for dimers 2 (spokes 3/4) and 4 (spokes 7/8). To improve the density for these map segments, we subjected the particles to another round of 3D-multibody refinement, in which the respective dimers were split into independently refined monomers. Using this approach, we achieved improvement of the local resolution for these density segments, which improved the resolution globally to 4.8 Å.

All resolution estimates were performed according to the gold standard FSC criterion of independently refined half maps (FSC = 0.143) within RELION. Local resolution was estimated using the RELION local post-processing implementation. The *B*-factor value used during local resolution filtering was –300.

Grouping of γ -TuRC spokes into structural clusters

First, GCPs were clustered according to cross-correlation between density segments. The γ -TuRC density was segmented into 14 spokes, comprising one GCP plus one γ -tubulin each. Pairwise cross-correlation between these segments was then computed in UCSF Chimera using the Fit in the Map command taking into account all density voxels.

Second, GCPs were clustered according to the r.m.s.d. values of docked atomic models. The models for human GCP4 (PDB code 3RIP) and γ -tubulin (PDB code 1Z5W) were docked into all 14 spokes of the γ -TuRC density. The GCP4 atomic model was split into four rigid bodies to account for interdomain conformational flexibility: Met1–Lys147, Ile148–Tyr361, Leu362–Lys505 and Ser506–Gln636. The very C-terminal helix Ile637–Tyr654 was manually moved into the density for the first spoke and then kept rigid with the last segment for the other positions. Docked GCP4 and γ -tubulin models were then combined into one PDB

file for each spoke and pairwise r.m.s.d. values were computed using the match command in UCSF Chimera. The average r.m.s.d. values computed between members of a specific cluster were at least two times smaller than the average r.m.s.d. values computed between members and nonmembers of a specific cluster (group (i): 0.75 Å versus 1.78 Å; group (ii): 0.60 Å versus 1.80 Å; group (iii) 1.61 Å versus 3.64 Å).

Identification of GCP-variant-specific structural features

Density segments covered by atomic models for GCP4 and γ -tubulin (prepared as described above) were annotated using the colour zone feature in UCSF Chimera. The remaining density segments for all spokes were compared against each other and correlated with primary sequence information on the different GCP variants guided by the docked GCP4 model, as described in the result section. Secondary structure prediction was performed in PSIPRED vs3.3^{31,32} and multiple sequence alignment was performed in PROMALS³³.

Atomic modelling

The crystal structure of human GCP4 (PDB code 3RIP) was used as template for homology modelling of *X. laevis* GCPs on the Phyre2 web portal³⁴ with one-to-one threading. The following UniProt sequences were used: XP_018080012.1 (GCP2), O73787 (GCP3), Q642S3 (GCP4), XP_018102626.1 (GCP5) and Q9DDA7 (GCP6, excluding the insertion domain sequence Leu532–Arg1260). The crystal structure of human γ -tubulin (PDB code 1Z5W) was used as template for homology modelling of *X. laevis* γ -tubulin with UniProt sequence P23330. The homology models were subsequently docked into the γ -TuRC density map using UCSF Chimera. Each model was split into several rigid bodies to account for interdomain conformational flexibility. The rigid bodies consisted of: Leu209–G364, Tyr365–Glu506, Glu507–Ser668, Ser671–Ser714, Ala715–Tyr867 for GCP2; Glu246–Gly389, Arg390–Tyr552, Asn553–Lys691, Gly692–Tyr885 for GCP3; Met1–Lys147, Ile148–Asp349, Ile350–Lys505, Ser506–Tyr654 for GCP4; Thr259–Pro421, Asp422–Leu723, Leu724–Asn847 and Lys848–Ala1014 for GCP5; Gln269–Ala421, Gly422–Leu1015, Lys1016–Ser1464, Asn1465–Tyr1622 for GCP6. The last three C-terminal helices had to be manually positioned for all homology models, because their arrangement differed significantly in the X-ray structure. Missing or incorrectly localized segments of the homology models were appended or adjusted in Coot³⁵, where the density allowed. This includes Leu215–Cys209 and Thr605–Gly611 in GCP2; Leu791–Glu815, Val818–Ile839, Glu246–Ser248 in GCP3; Thr259–Gln265 and Arg571–Leu582 after removal of Arg571–Ser670 in GCP5; Gln269–Asp279 and Ser556–Met561 after removal of Phe532–Leu557 in GCP6. Unassigned, but clearly resolved, α -helices in the belt density were built as poly-alanine helices in Coot. In total, 17 helices were built with the following number of residues: 16, 19, 15, 13, 13, 17, 14, 15, 16, 16, 14, 13, 23, 15, 18, 13 and 31. Parts of the homology models not resolved in the cryo-EM reconstruction were removed. Rigid-body-fitted homology models were combined in Chimera into one model and main chain breaks caused by domainwise rigid-body fitting were repaired in Coot. The model was refined in real space against the cryo-EM density and structural clashes were removed using molecular dynamics flexible fitting (MDFF³⁶). MDFF simulations were prepared using QwikMD³⁷ and carried out with NAMD³⁸ using the CHARMM36 force field. Secondary structure, *cis* peptide and chirality restraints were used during 800 steps of minimization followed by a 40-ps simulation at 300 K. The resulting model was refined in Phenix 1.14 using the initial model as reference and then submitted to the NAMDinator website tool³⁹ that combines molecular dynamics flexible fitting (MDFF) with real-space refinement in Phenix. MDFF was performed at 298 K using 2,000 minimization steps and 20,000 simulation steps. Simulation was run in vacuo with scaling factor of 0.3. Subsequently, the fit of resolved bulky amino acid side chains was inspected and corrected in Coot, if required, before the model was refined in Phenix. This procedure was repeated twice with two slightly different refinement parameters to

obtain the final model. Model validation was performed in Phenix 1.14⁴⁰. The register of the atomic model was confirmed by the following list of resolved bulky amino acid side chains (spoke number/residue/unique as defined in Extended Data Fig. 5): 1/His359/-, 1/Tyr377/-, 1/Tyr455/*, 1/His725/-, 3/Phe275/, 3/Trp324/*, 3/Tyr455/*, 3/His724/-, 5/Trp324/*, 5/Phe325/*, 5/Arg486/-, 5/Tyr493/-, 5/Phe501/-, 5/His643/-, 5/Tyr646/*, 5/His649/-, 5/His724/-, 7/His289/-, 7/Trp324/*, 7/His359/*, 7/Phe363/*, 7/Tyr455/*, 7/Tyr491/-, 7/His511/-, 7/Trp705/-, 13/Phe275/-, 13/Trp324/*, 13/His359/-, 2/Trp298/-, 2/His386/-, 2/His658/*, 2/Phe665/-, 4/Trp298/-, 4/Phe369/*, 4/Phe371/*, 4/His416/-, 4/Trp667/-, 6/Trp298/-, 6/His342/-, 6/Phe369/-, 6/Trp371/-, 6/Tyr423/*, 6/Phe491/*, 6/Phe529/-, 6/Phe538/-, 6/Phe665/-, 6/Trp680/-, 10/Phe371/*, 10/His409/*, 10/Tyr437/-, 10/Phe458/*, 10/Trp461/*, 10/Tyr703/-, 10/Tyr816/-, 10/Phe820/-, 8/Trp298/-, 8/Tyr306/-, 8/Phe323/-, 8/Tyr334/*, 8/Tyr335/*, 8/His342/-, 8/Phe371/-, 8/His386/-, 8/His416/-, 8/Tyr423/*, 8/Phe529/-, 8/Tyr543/*, 8/His557/-, 8/Tyr719/-, 8/His750/-, 8/His851/-, 14/Trp298/-, 14/Tyr402/-, 14/His493/-, 14/Tyr537/-, 14/Phe538/-, 14/Phe852/-, 9/Lys47/*, 9/Arg54/-, 9/Phe55/*, 9/Phe58/*, 9/Tyr124/-, 9/Phe129/-, 9/Tyr158/-, 9/Tyr184/-, 9/His390/-, 9/His401/-, 9/Phe478/-, 9/His560/-, 9/His562/*, 9/His580/*, 9/His590/*, 11/Phe58/*, 11/His121/*, 11/His130/*, 11/Tyr158/-, 11/Tyr184/-, 11/Trp190/-, 10/Trp272/-, 10/Tyr316/-, 10/Trp362/*, 10/Phe368/-, 10/His925/-, 10/His933/*, 12/His340/-, 12/Tyr367/-, 12/Tyr400/*, 12/Tyr431/*, 12/Tyr433/*.

Unbiased structure-guided identification approach

For the list of proteins identified in the purified γ -TuRC by LFQ mass spectrometry, we either downloaded X-ray structures from the PDB or—where possible—prepared high-confidence homology models in HHpred. Subsequently, the fitmap command of UCSF Chimera was used to fit these atomic models into the envelope of the cryo-EM density, starting from 10,000 randomly sampled starting positions and orientations for each model. The following parameters were used: correlation metric with simulated densities at 6 Å resolution, 50 Å radius around the starting position, global search. For each relevant model, the number of fits was plotted against the respective cross-correlation coefficients (Extended Data Fig. 6). Atomic models for NEDD1, NME7 and MOZART1 were limited to protein segments with high homology to known structures (1–361 for NEDD1; 26–58 for MOZART1; 91–226 and 235–373 for NME7). The following list of proteins could not be localized in the γ -TuRC density even though highly abundant in the γ -TuRC purification according to LFQ mass spectrometry (UniProt code/protein name/template probability from homology modeling/LFQ intensity): AOA1L8H345/dynamin-1/100/8.72E+09, Q8AVE2/Hsc70 protein/100/5.89E+09, AOA1L8FKY3/Hsp70/100/5.5E+09, AOA1L8HW84/tight junction protein ZO-3/100/5.47E+09, Q7ZTN1/tight junction protein ZO-3/100/5.42E+09, AOA1L8GWY3/protein transport protein Sec24a/100/5.26E+09, AOA1L8FW10/insulin-like growth factor 2 mRNA-binding protein 3/99.95/4.34E+09, AOA1L8EKZ2/polyadenylate-binding protein/100/4.32E+09, AOA1L8GIU5/protein transport protein SEC23/100/3.92E+09, AOA1L8ES55/polyadenylate-binding protein/100/3.72E+09, AOA1L8H4P1/CSD_1 domain-containing protein/99.35/3.36E+09, AOA1L8G7U0/protein transport protein SEC23/100/2.75E+09, AOA1L8GRB6/uncharacterized protein/100/2.47E+09, AOA1L8GY92/WD_REPEATS_REGION domain-containing protein/99.96/1.77E+09, AOA1L8HEX9/uncharacterized protein/100/1.67E+09, AOA1L8FA78/heat shock-related 70 kDa protein; signalling protein/100/1.62E+09, AOA1L8FIH5/dynamin-1/100/1.47E+09, AOA1L8HWC1/uncharacterized protein/100/1.45E+09, AOA1L8GMZ9/dynamin-1/100/1.39E+09, Q8AVK9/NSEPI protein/99.33/1.15E+09, AOA1L8F457/protein transport protein SEC23/100/1.07E+09, AOA1L8FAZ8/protein transport protein SEC23/100/1.01E+09, AOA1L8GWQ5/WD repeats region domain-containing protein/100/8.22E+08, AOA1L8HF79/nucleoside diphosphate kinase 7/99.92/7.27E+08, AOA1L8G3Y8/fragile X mental retardation syndrome-related protein 1/100/7.13E+08, A3KMH8/fragile X mental

retardation syndrome-related protein 1/100/7.11E+08, AOA1L8EWC9/interferon-inducible double-stranded RNA-dependent protein kinase activator A/100/6.98E+08, AOA1L8FKW5/serine/threonine-protein kinase TOR/100/6.7E+08, AOA1L8F6I3/HSPA5 protein/100/6.64E+08, AOA1L8GQQ7/uncharacterized protein/100/5.84E+08, AOA1L8HM56/protein transport protein Sec24B/100/5.23E+08, AOA1L8FTJ1/polyadenylate-binding protein/100/4.81E+08, AOA1L8FZR3/polyadenylate-binding protein/100/4.81E+08, AOA1L8EM44/ γ -actin/100/4.77E+08, AOA1L8ETE5/uncharacterized protein/100/4.77E+08, AOA1L8HRT0/DZF domain-containing protein/100/4.46E+08, AOA1L8GT63/tight junction protein ZO-1/100/4.42E+08, Q6GMC1/ubiquitin-40S ribosomal protein S27a/100/4.27E+08, AOA1L8GSV1/ubiquitin-like domain-containing protein/100/4.27E+08, AOA1L8HQK4/polyubiquitin-C/100/4.26E+08, Q7SY79/Ubc-prov protein/100/4.26E+08, Q6GQF3/ubiquitin-60S ribosomal protein L40/100/4.26E+08, AOA1L8HX68/ubiquitin-like domain-containing protein/99.94/4.26E+08, AOA1L8HCZ9/uncharacterized protein/100/4.26E+08, AOA1L8H6E1/nucleoside diphosphate kinase 7/99.92/4.13E+08.

Analysis of geometrical parameters

For analysis of geometrical parameters of the γ -TuRC, we first computed the (approximate) helical axis for each of the analysed models by fitting a straight line through the centroids described below. For the γ -TuRC, we defined centroids for five different sets of atoms along the spokes (Fig. 4a) using UCSF Chimera. Two points were defined on the γ -tubulin subunit of each spoke (γ -tubulin Thr145, γ -tubulin Tyr152), while the other three points were defined based on aligned amino acids of the GCPs (point 1: Leu554 of GCP3, Leu508 of GCP2, Met590 of GCP6, Leu350 of GCP4, Leu711 of GCP5; point 2: Pro408 of GCP3, Ser369 of GCP2, Tyr445 of GCP6, Pro166 of GCP4, Glu451 of GCP5; point 3: Leu249 of GCP3, Leu216 of GCP2, Met1 of GCP4, Val266 of GCP5, Leu280 of GCP6). For the 13-spoked MT (based on PDB 6EVZ), three centroids were defined based on Gln15 of three different layers in the MT. For the yeast γ -TuSC in the closed conformation (PDB 5FLZ), three centroids were defined based on yeast γ -tubulin Gln12, Asn389 of GCP2/Asn444 of GCP3 and Phe102 of GCP2/Phe216 of GCP3. All three models were subsequently aligned according to their central axis and coordinates were transformed such that the axes correspond to the z-axis of the coordinate system. Cartesian coordinates of tubulins were determined based on conserved residues (Gln15, Gln16 and His16). On the basis of these coordinates, we computed the pitch (Fig. 4b) for each spoke and radial distances (Fig. 4c).

Using UCSF Chimera, refined homology models for GCPs and γ -tubulins (see above) were docked as rigid bodies into the cryo-EM density of the yeast γ -TuSC oligomer in a closed conformation (EMD-2799). The atomic models of the *X. laevis* γ -TuRC in the experimentally observed conformation and the simulated closed conformation were aligned with respect to spoke 1 using UCSF Chimera. The conformational changes linking both models were captured in motion by interpolating between the two conformations using UCSF Chimera's morph conformations function. In an alternative approach to visualize the conformational change, we used PyMol⁴¹ to draw vectors between C α atoms in the two conformations and colour-coded them according to their r.m.s.d. values.

To analyse relative GRIP1–GRIP2 domain inclinations, we computed their axes for each spoke. To define the GRIP1 domain axis, we used the following corresponding residue pairs: Ala266/Gly453 for GCP2, Gly297/Gly486 for GCP3, Gly49/Gly280 for GCP4, Gly317/Gly559 for GCP5 and Gly332/Gly526 for GCP6. To define the GRIP2 domain axis, we used the following corresponding residue pairs: Leu508/Phe757 for GCP2, Leu554/Phe784 for GCP3, Leu350/Phe592 for GCP4, Leu711/Phe959 for GCP5 and Met1319/Phe1553 for GCP6. The relative inclinations of GRIP1 and GRIP2 domains were measured in UCSF Chimera with the angle command. The GRIP1–GRIP2 inclinations were computed separately for each spoke and then quantified for each GCP

variant: five copies of GCP2, five copies of GCP3, two copies of GCP4, one copy of GCP5 and one copy of GCP6. The average angle for each GCP group was calculated as mean \pm s.e.m.

To analyse interspoke distances, we computed distances between conserved residues in the GCP GRIP1 and GRIP2 domains or γ -tubulins of neighbouring spokes in UCSF Chimera. The following residues were used: Ala266 of GCP2, Gly297 of GCP3, Gly49 of GCP4, Gly317 of GCP5 and Gly332 of GCP6 for the GRIP1 domain; Phe757 of GCP2, Phe784 of GCP3, Phe592 of GCP4, Phe959 of GCP5 and Phe1553 of GCP6 for the GRIP2 domain; Asn187 for γ -tubulin.

Research animals

The permission numbers for *X. laevis* experiment are: 35-9185.81/G-204/12 (Regierungspräsidium Karlsruhe, BW); 81-02.05.40.17.091 (LANUF Recklinghausen, NRW). Ethic justification of the experiment is according to § 7a Abs. 2 Nr. 3 TierSchG. We have complied with all relevant ethical regulations. The injection of hormones into frogs mimics the physiological stimulus upon reproduction with a negligible burden for animals. In turn, the experiments in *X. laevis* using laid eggs and derived extracts have contributed ground-breaking knowledge to basic mechanisms of cell division, cell cycle control and MT functions. The latter has led, for example, to the discovery and characterization of cytostatic drugs such as taxol (Paclitaxel). These experiments are therefore performed with an optimal cost–benefit ratio, and show hardly any harm to animals while allowing great potential for gaining knowledge for basic and applied science.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Cryo-EM densities of the γ -TuRC filtered according to global or local resolution have been deposited in the Electron Microscopy Data Bank (EMDB) under accession code EMD-10491. Atomic coordinates for the γ -TuRC have been deposited at the PDB under accession code 6TF9. The original immunoblots and further source data from LFQ mass spectrometry (Fig. 2a), immunoblot quantification (Fig. 3c), geometric analysis of the atomic model (Figs. 3g, 4b–d), MT nucleation assays (Extended Data Figs. 1c, 6i, 9b, d), circular dichroism measurements (Extended Data Fig. 4e), unbiased structure-guided identification (Extended Data Fig. 6a–d), quantification of indirect immunofluorescence (Extended Data Fig. 6g) and actin polymerization (Extended Data Fig. 6h) are included in the Supplementary Information. The raw cryo-EM micrograph movie stacks are available from the corresponding authors upon request.

16. Zhang, L., Keating, T. J., Wilde, A., Borisy, G. G. & Zheng, Y. The role of Xgrip210 in γ -tubulin ring complex assembly and centrosome recruitment. *J. Cell Biol.* **151**, 1525–1536 (2000).
17. Chinen, T. et al. The γ -tubulin-specific inhibitor gatastatin reveals temporal requirements of microtubule nucleation during the cell cycle. *Nat. Commun.* **6**, 8722 (2015).
18. Bärenz, F. et al. The centriolar satellite protein SSX2IP promotes centrosome maturation. *J. Cell Biol.* **202**, 81–95 (2013).
19. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
20. Schwanhäusser, B. et al. Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 (2011).
21. Tang, G. et al. EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
22. van Heel, M., Harauz, G., Orlova, E. V., Schmidt, R. & Schatz, M. A new generation of the IMAGIC image processing system. *J. Struct. Biol.* **116**, 17–24 (1996).
23. Liu, X. & Wang, H. W. Single particle electron microscopy reconstruction of the exosome complex using the random conical tilt method. *J. Vis. Exp.* **49**, 2574 (2011).

24. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
25. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
26. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
27. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
28. Zivanov, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCr* **6**, 5–17 (2019).
29. Nakane, T., Kimanius, D., Lindahl, E. & Scheres, S. H. W. Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *eLife* **7**, e36861 (2018).
30. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
31. Buchan, D. W. A. & Jones, D. T. The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Res.* **47**, W402–W407 (2019).
32. Jones, D. T. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292**, 195–202 (1999).
33. Pei, J., Tang, M. & Grishin, N. V. PROMALS3D web server for accurate multiple protein sequence and structure alignments. *Nucleic Acids Res.* **36**, W30–W34 (2008).
34. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protocols* **10**, 845–858 (2015).
35. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
36. Trabuco, L. G., Villa, E., Schreiner, E., Harrison, C. B. & Schulten, K. Molecular dynamics flexible fitting: a practical guide to combine cryo-electron microscopy and X-ray crystallography. *Methods* **49**, 174–180 (2009).
37. Ribeiro, J. V. et al. QwikMD — integrative molecular dynamics toolkit for novices and experts. *Sci. Rep.* **6**, 26536 (2016).
38. Phillips, J. C. et al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
39. Kidmose, R. T. et al. NAMDinator - automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. *IUCr* **6**, 526–531 (2019).
40. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
41. Delano, W. L. PyMOL: an open-source molecular graphics tool http://www.ccp4.ac.uk/newsletters/newsletter40/11_pymol.pdf (2018).
42. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
43. Sievers, F. & Higgins, D. G. Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci.* **27**, 135–145 (2018).
44. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).

Acknowledgements We thank A. Boland for graphene oxide grids, G. Hofhaus and T. Mielke for technical cryo-EM support, S. Hata for suggestions, S. Kmiecik for assisting with circular dichroism measurements, L. Rohland for assisting with circular dichroism and actin polymerization measurements. We acknowledge the services SDS@hd and bwHPC supported by the Ministry of Science, Research and the Arts Baden-Württemberg, as well as the German Research Foundation (INST 35/1314-1 FUGG and INST 35/1134-1 FUGG). We also acknowledge support from the German Research Foundation to E.S. (Schi 295/4-3), O.J.G. (GR-1737/9-1) and T. Ruppert (Ru 747/1-1). P.L. received a HBIGS fellowship.

Author contributions P.L. optimized γ -TuRC purification, performed γ -TuRC activity measurements, salt treatments, CEP215-N analyses in vitro and in egg extracts, GCP6 insertion domain alignment, sucrose gradient and silver staining. P.L. and A.N. performed negative-stain electron microscopy and particle selection. E.S. and P.L. performed the DNaseI inhibition experiment. D.F. performed 2D classification of negatively stained γ -TuRC. J.L. and C.S. performed initial cryo-EM analysis of the γ -TuRC. P.L. and A.N. prepared, optimized and screened γ -TuRC cryo-EM grids and, together with E.Z., selected particles for initial model generation. P.L. performed immunoprecipitation experiments. A.B. performed circular dichroism measurements and actin polymerization assay. T. Ruppert performed LFQ mass spectrometry analysis. A.N. and A.B. detected actin in the γ -TuRC by immunofluorescence. C.P. prepared egg extracts. O.J.G. provided egg extracts and advised P.L. with the γ -TuRC purification. E.Z. analysed the negative-staining data of γ -TuRC and CEP215-N, acquired and analysed cryo-EM data, and, together with T. Rudack, prepared atomic models and pursued all related aspects of structure analysis. All authors discussed and interpreted the γ -TuRC cryo-EM density. E.S. and S.P. planned and supervised the experiments, analysed data and, together with E.Z. and P.L., wrote the manuscript.

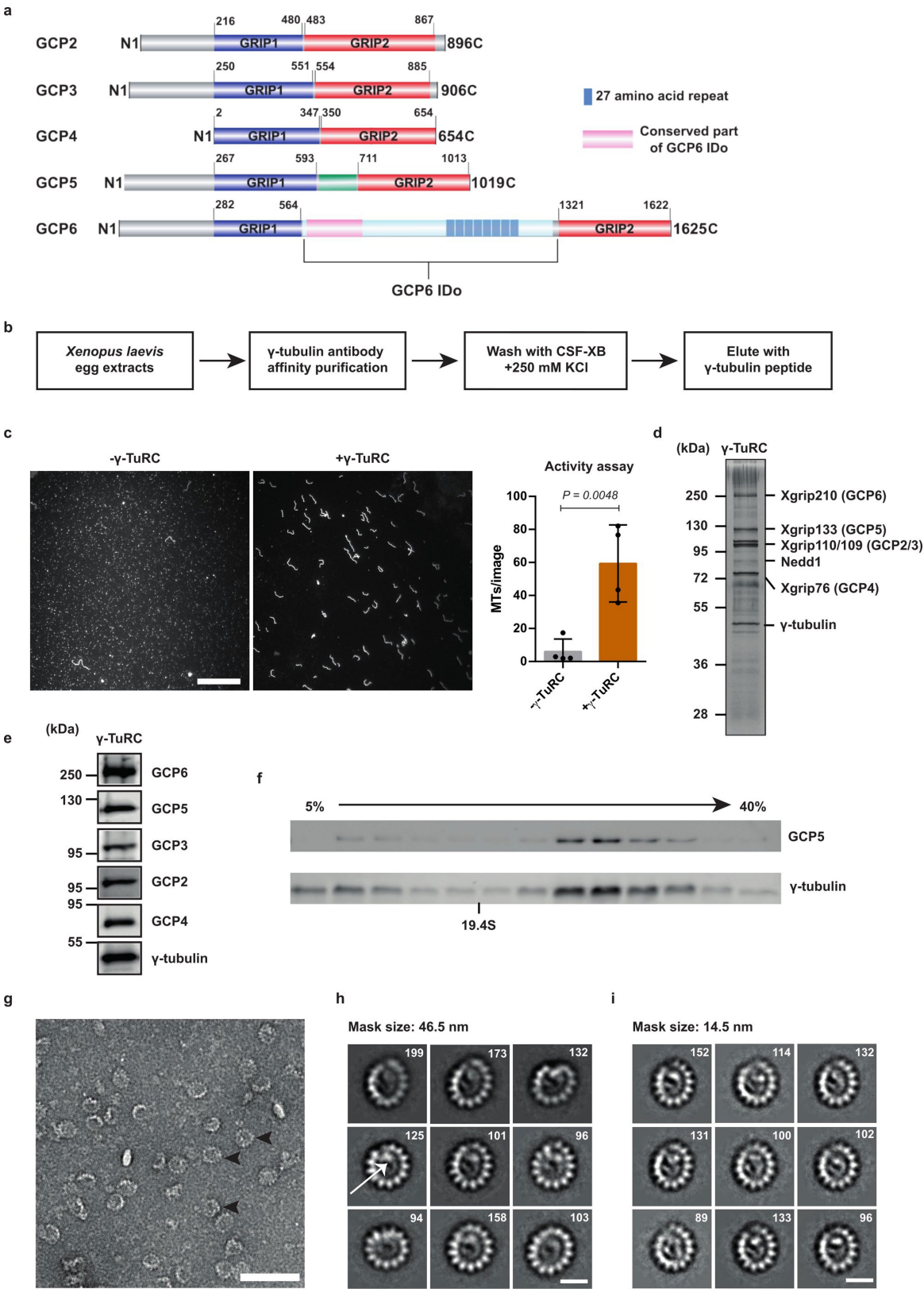
Competing interests The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-019-1896-6>.

Correspondence and requests for materials should be addressed to S.P. or E.S.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



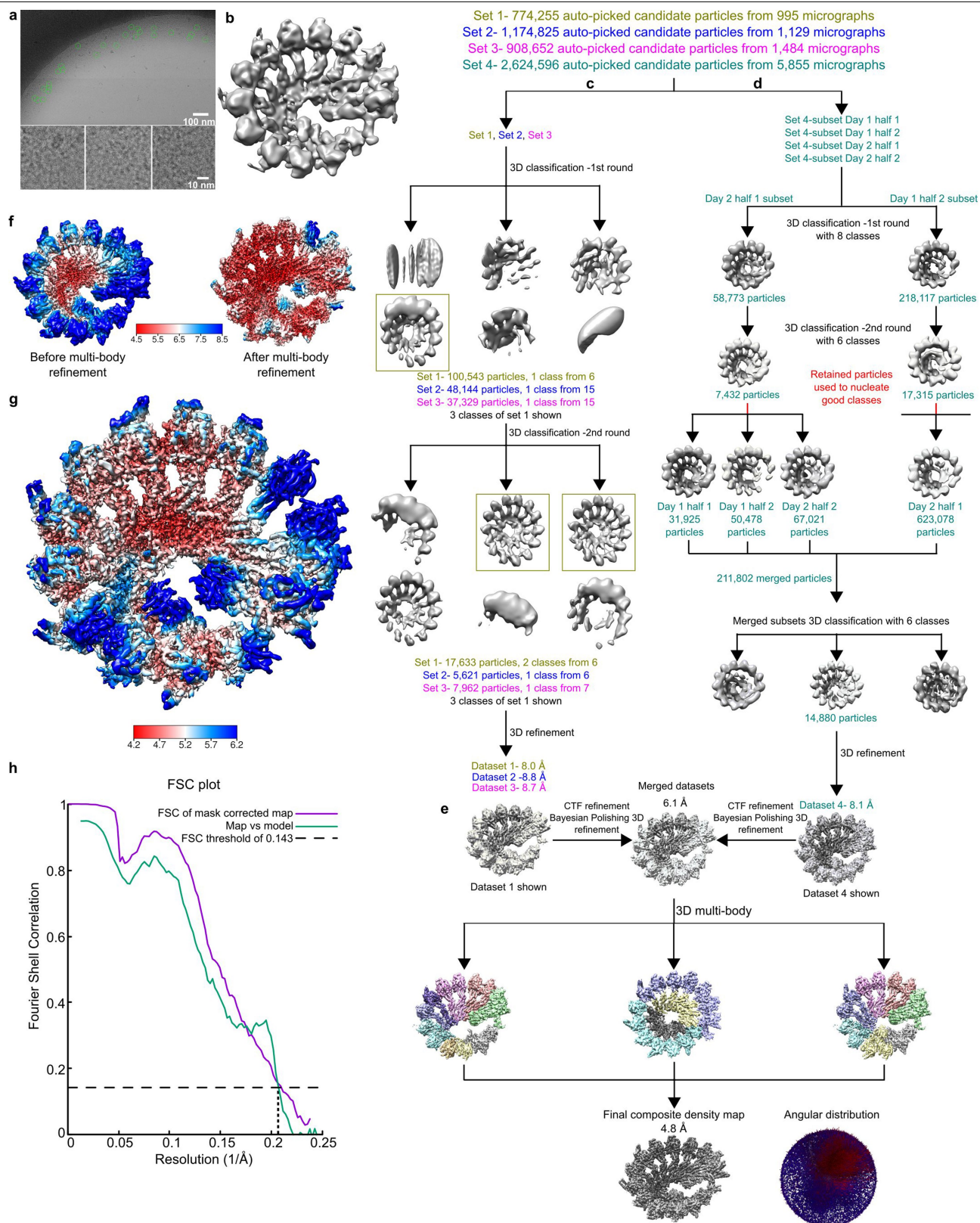
Extended Data Fig. 1|See next page for caption.

Extended Data Fig. 1 | Biochemical, functional and structural

characterization of γ -TuRCs purified from *X. laevis* egg extracts. **a**, Domain organization of GCP2–GCP6. Between the conserved GRIP1 and GRIP2 domains, GCP5 and GCP6 possess a 120- and 750-residue-long insertion domain, respectively. The GCP6 insertion domain contains 8 repeats of 27 amino acids. Domains are annotated according to the Pfam database.

b, Schematic of γ -TuRC purification. The γ -TuRC was purified with γ -tubulin antibody-crosslinked Protein A Dynabeads, washed with CSF-XB buffer containing 250 mM salt (KCl), and then eluted by a short peptide corresponding to the C terminus of γ -tubulin. **c**, Purified γ -TuRCs showed basal MT nucleation activity. Experiment was carried out with (+) or without (–) purified γ -TuRCs, and 33 μ M tubulin (5% Cy3-labelled tubulin for visualization). In the negative control, the same purification procedure was used with eluates from rabbit random IgG-crosslinked Protein A Dynabeads. Twenty random images were acquired with a light microscope, and representative overview images are shown. Right, the number of MTs was quantified by ImageJ and data are mean \pm s.d. $n = 4$ biologically independent experiments. P value determined by unpaired two-sided t -test. Scale bar, 10 μ m. **d, e**, After the affinity purification of γ -TuRCs, the eluted proteins were resolved by SDS–PAGE followed by silver staining (**d**) and immunoblotting (**e**). Representative images

in **d** and **e** are from three biologically independent experiments. For gel source data, see Supplementary Fig. 1. **f**, Immunoblotting analysis of the purified γ -TuRCs after sucrose gradient. Purified γ -TuRCs were applied to a 5–40% sucrose gradient and fractionated after centrifugation. Fractions were resolved by SDS–PAGE and probed using γ -tubulin and GCP5 antibodies. Thyroglobulin (19.4 S) was used as a standard marker and run on a parallel gradient. Representative images were from three biologically independent experiments. For gel source data, see Supplementary Fig. 1. **g**, Confirmation of structural integrity of the purified γ -TuRCs by negative-staining electron microscopy. Representative micrograph is from five biologically independent experiments. Scale bar, 100 nm. Black arrowheads denote examples of particles used in 2D classification and averaging. **h, i**, γ -TuRC particles from the negative-stain electron microscopy were classified and averaged with a mask size of either 46.5 nm (**h**) or 14.5 nm (**i**). Representative classes of γ -TuRCs are from three biologically independent experiments. The number of particles contributing to each class is given. An example of the ‘asymmetric’ density inside the γ -TuRC is highlighted by a white arrow (**h**), which is more readily visible with a 14.5-nm mask focusing on the inner part of the γ -TuRC (**i**). Scale bars, 20 nm.

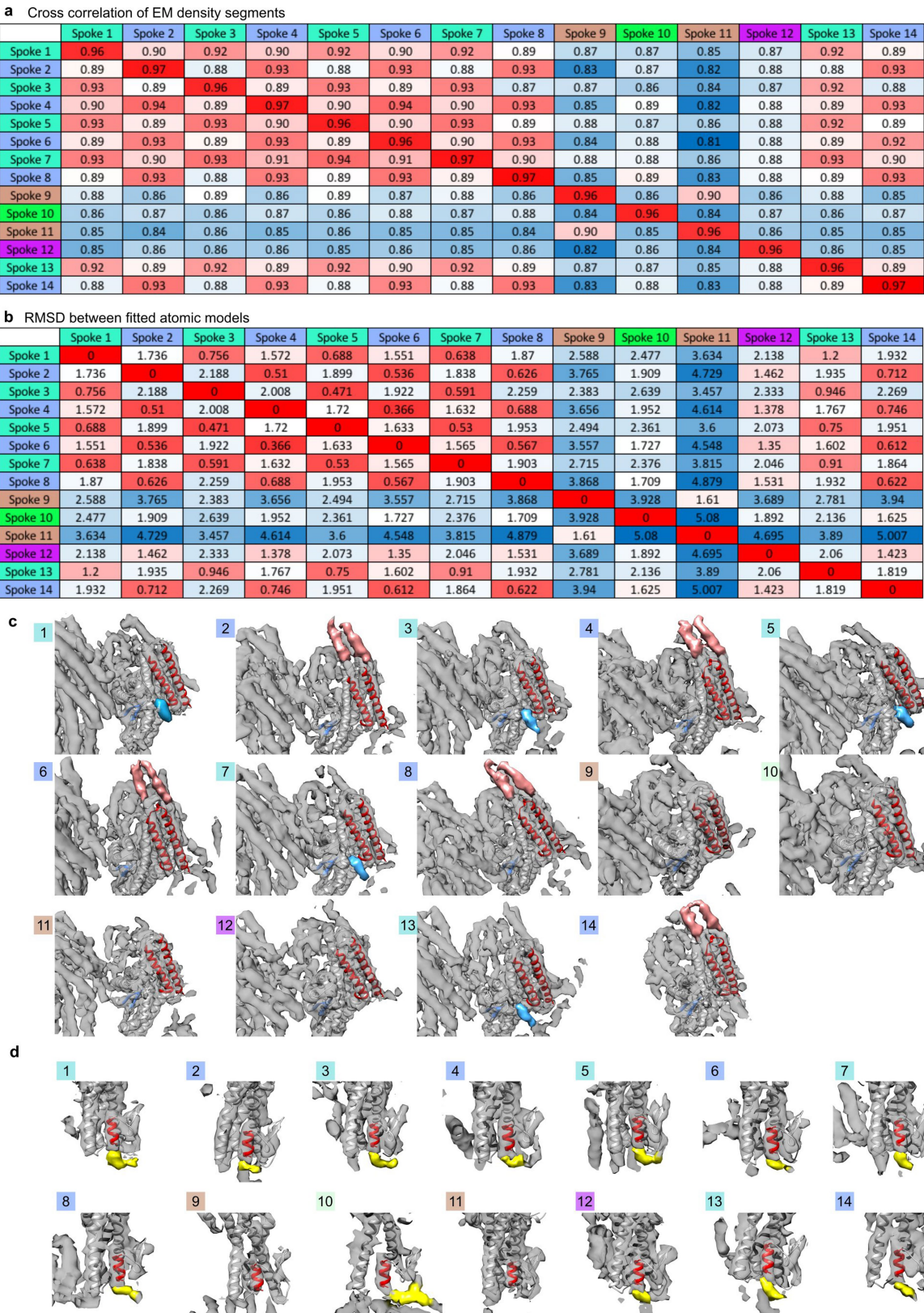


Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Cryo-EM data processing and resolution estimation.

a, A representative micrograph from four biologically independent experiments with manually selected particles (green circles) is shown. Three selected particles are shown below. **b**, Initial model of the γ -TuRC obtained from 3,000 manually selected particles (Methods). Scale bars are depicted in **a**. **c**, Four datasets were acquired and initially processed separately. Datasets 1, 2 and 3 were submitted to two consecutive rounds of 3D classification with a varying number of classes. All class averages obtained for dataset 1 are shown. Retained class averages are highlighted by a rectangular box. **d**, Dataset 4 was the largest dataset, and was therefore divided into four subsets. The higher-quality subsets were submitted to two rounds of 3D classification and particles encompassed in the high-quality classes from these two subsets were combined with all four original subsets of particles to nucleate high-quality classes in the two lower-quality subsets. Only class averages retained for further processing are shown. **e**, The final sets of particles from all four

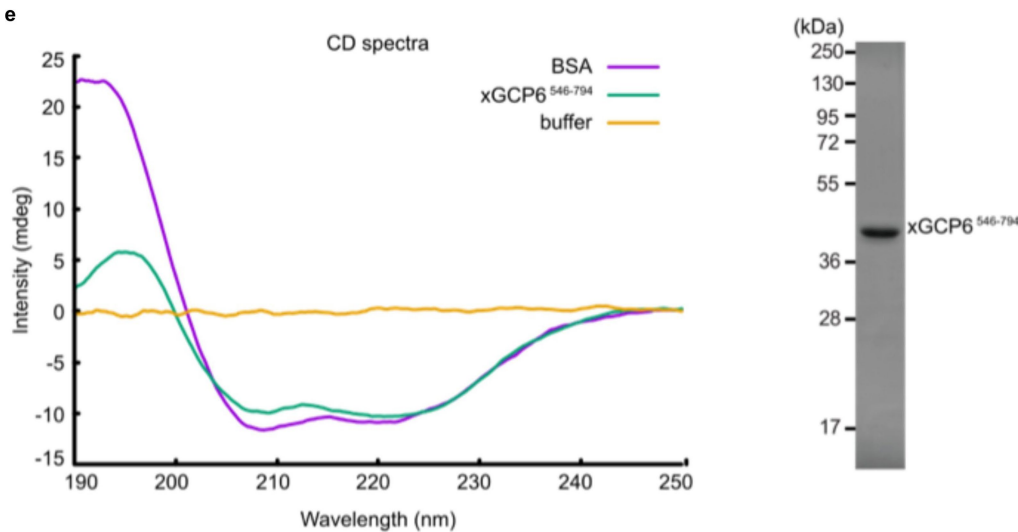
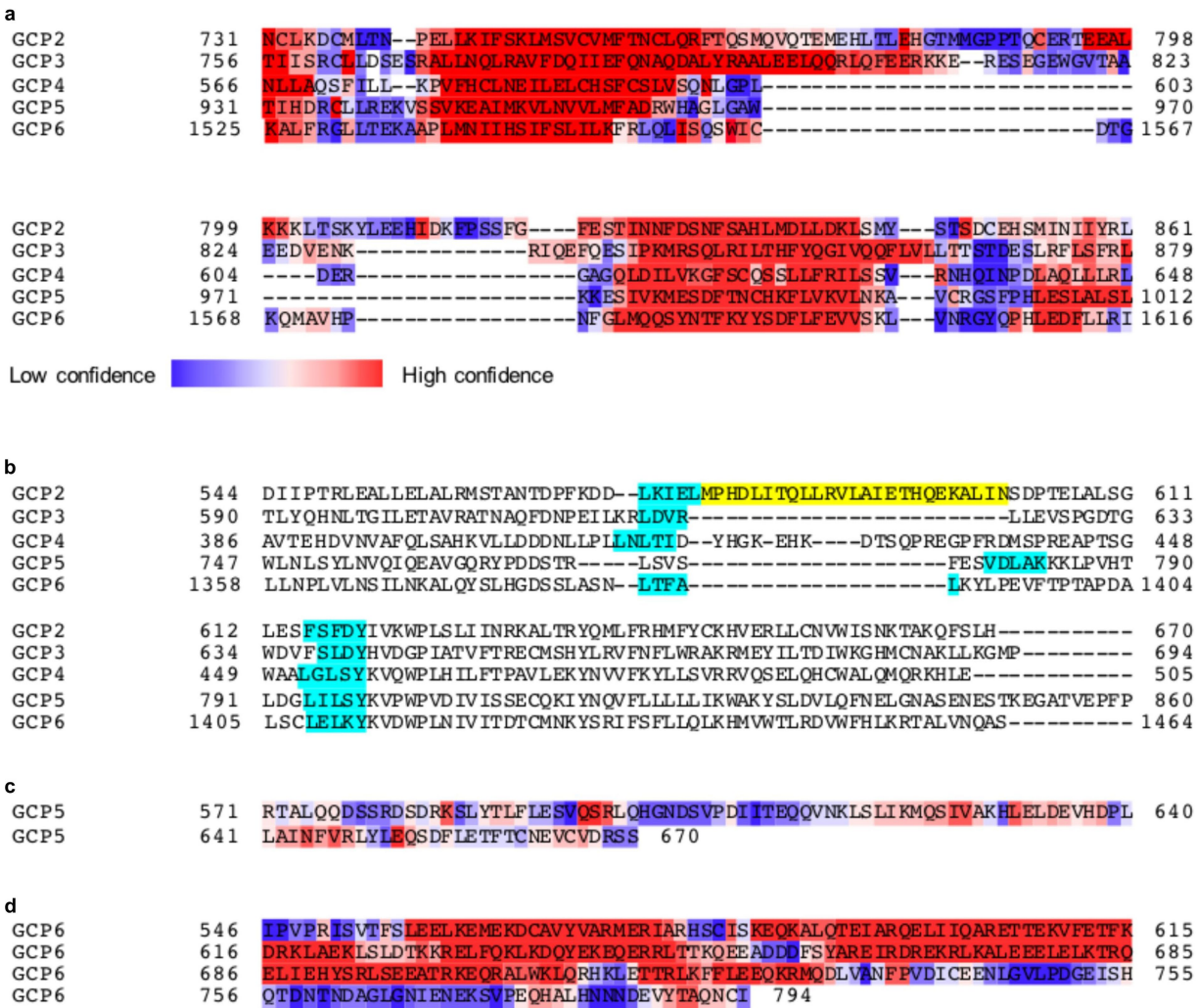
datasets were refined separately, submitted to CTF refinement and Bayesian polishing, and subsequently merged. The γ -TuRC density was split into several combinations of segments, as indicated. All combinations of segments were subjected to a multibody 3D refinement separately. Their output density segments were combined into one composite density for analysis. The angular distribution of particle views is shown. **f**, Comparison of local resolution estimation before (left) and after (right) 3D multibody refinement. Local resolution was markedly improved for the peripheral segments of the γ -TuRC density reconstruction upon 3D multibody refinement. The colour-coded resolution scale is identical for both panels. **g**, Local resolution estimation of the final γ -TuRC density map with a resolution range covering the entire spectrum. **h**, Mask-corrected FSC between the two independently refined half-set reconstructions (purple), and between the full reconstruction and the atomic model for the γ -TuRC (green).



Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Structural grouping and unique GCP-variant-specific features identify GCP proteins in the γ -TuRC. **a**, Pairwise cross-correlation between isolated density segments, colour-coded from higher (red) to lower (blue) correlation. The actual correlation values are given. **b**, Atomic models of human GCP4 and γ -tubulin were fitted into the 14 γ -TuRC spokes domainwise (Methods). Pairwise r.m.s.d. between C α atoms of atomic models representing the individual spokes, colour-coded from lower (red) to higher (blue) r.m.s.d. values. The actual r.m.s.d. values are given. Both approaches cluster the spokes into five classes, colour-coded in the left column and top row. **c**, **d**, Atomic models for human GCP4 (white) were fitted domainwise into the density.

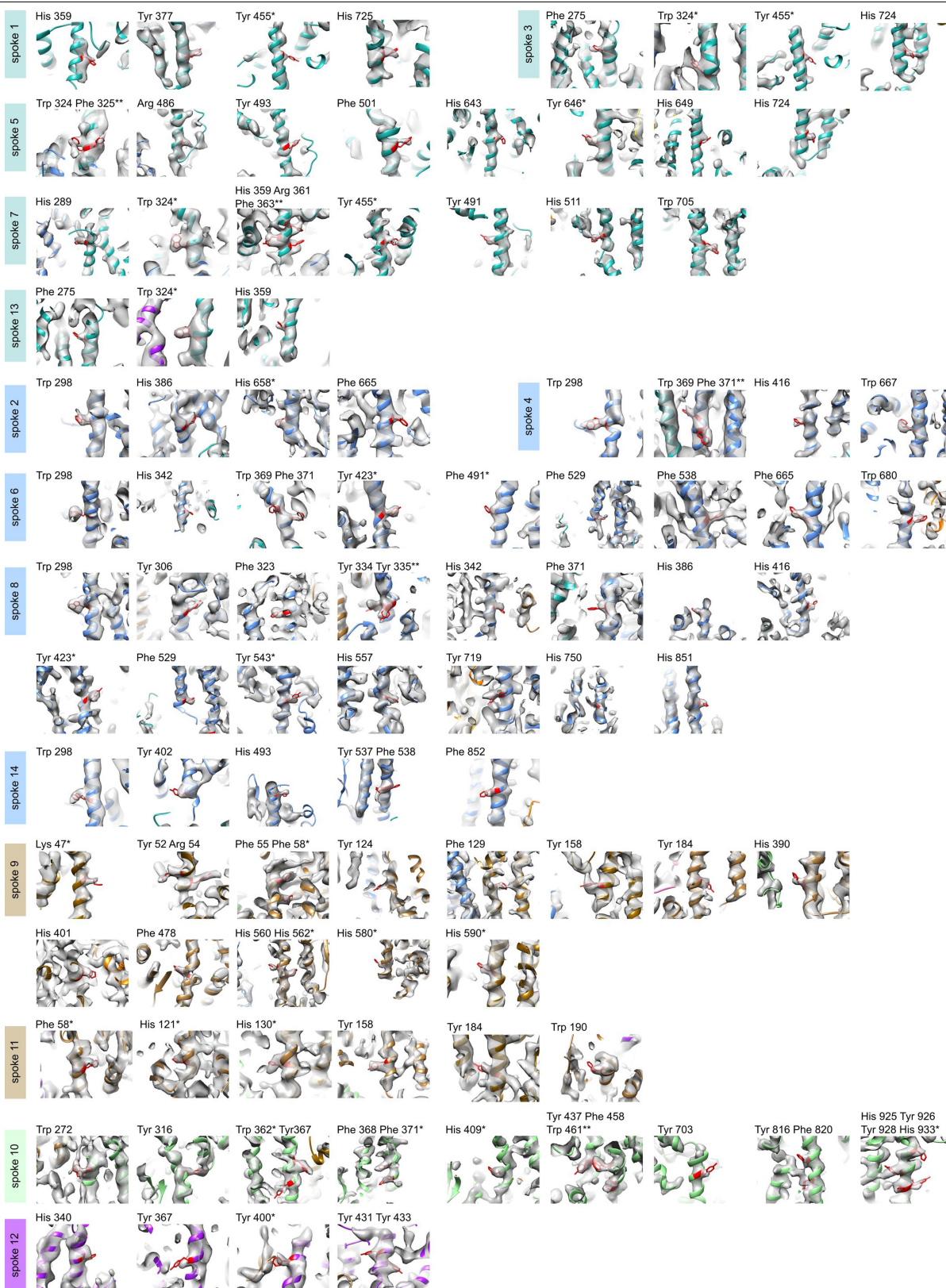
Density segments covered by the atomic model are depicted in transparent grey. Remaining unexplained segments are depicted in colour. Features are shown for all 14 individual spokes of the γ -TuRC. **c**, Characteristic density segments of the GRIP2 domain. Extended C-terminal α -helices (red model and density) are unique for group (ii) (spokes 2, 4, 6, 8, 14), and an extended loop between the GRIP2 β -strands (blue model and density) is present only in group (i) (spokes 1, 3, 5, 7, 13). **d**, Unexplained density segments N-terminal of the GRIP1 domain. Only group (iii) (spokes 9, 11) is devoid of a continuous density connecting to the N-terminal helix of GCP4 (position 9 and 11; lack of yellow extension). Colour code as in Fig. 1a.



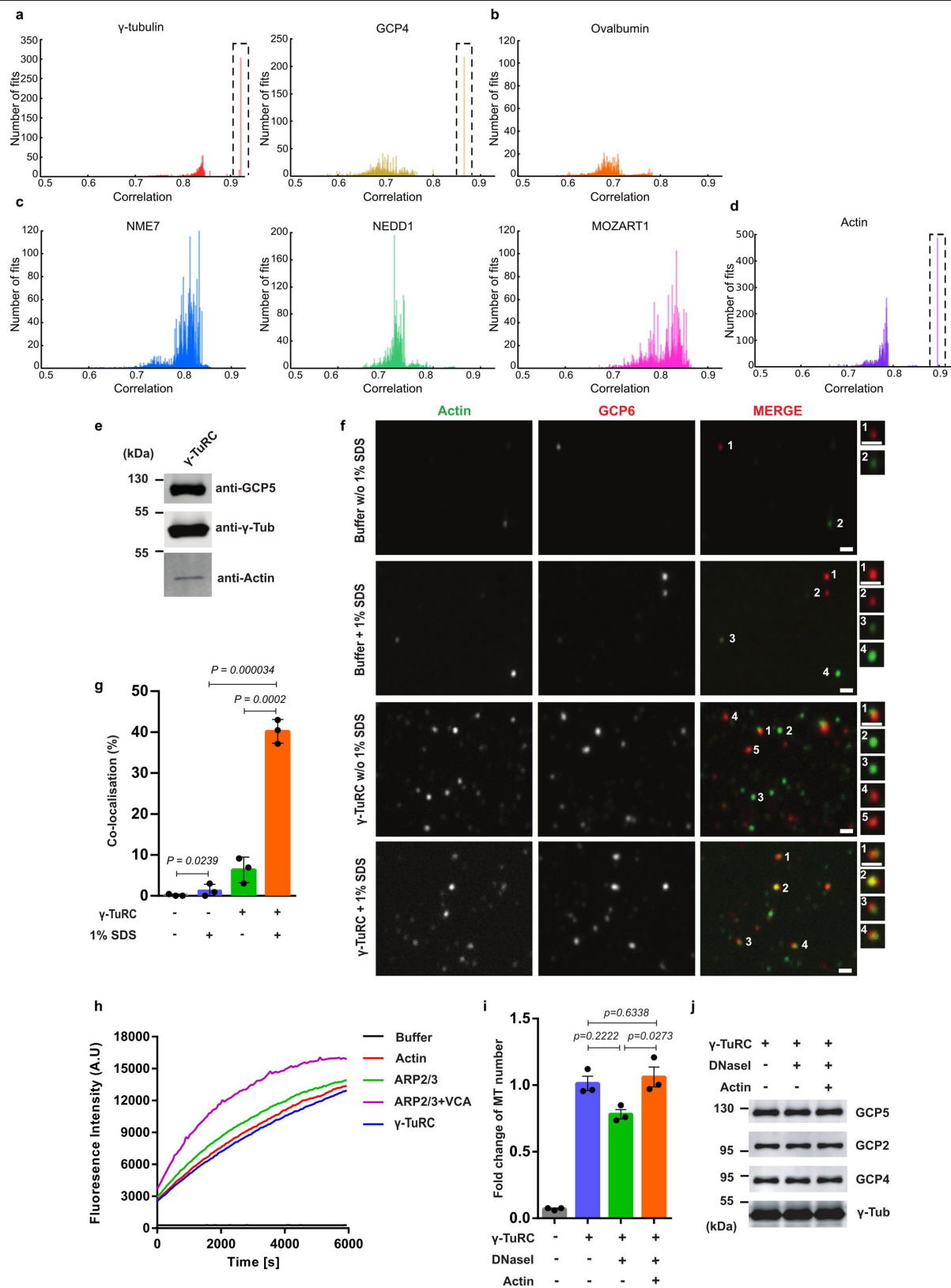
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Secondary structure prediction and biochemical characterization for GCP variant-distinguishing regions and the GCP5 and GCP6 insertion domains. **a**, C-terminal segments of all GCP variants were aligned and the secondary structure was predicted. Confidence for α -helical secondary structure is colour-coded (blue denotes low confidence, red denotes high confidence). **b**, Multiple sequence alignment for the GRIP2 segments encompassing the inter- β -strand loop for all GCP variants. β -strands are highlighted in blue. **c**, Secondary structure prediction for the insertion domain of GCP5. The prediction confidence for α -helical secondary structure is

colour-coded. **d**, Secondary structure prediction of xGCP6(546–794) showing highly α -helical character. **e**, Circular dichroism (CD) analysis of the purified xGCP6(546–794). Left, representative plots are from three biologically independent experiments. Comparison of the circular dichroism spectra for xGCP6(546–794) and BSA (containing only α -helices) confirms the predicted α -helical character of the N-terminal part of the GCP6 insertion domain. Right, Coomassie-blue-stained SDS–PAGE gel of the elution fraction of the xGCP6(546–794) shows the purity of the sample. Representative image is from three biologically independent experiments.



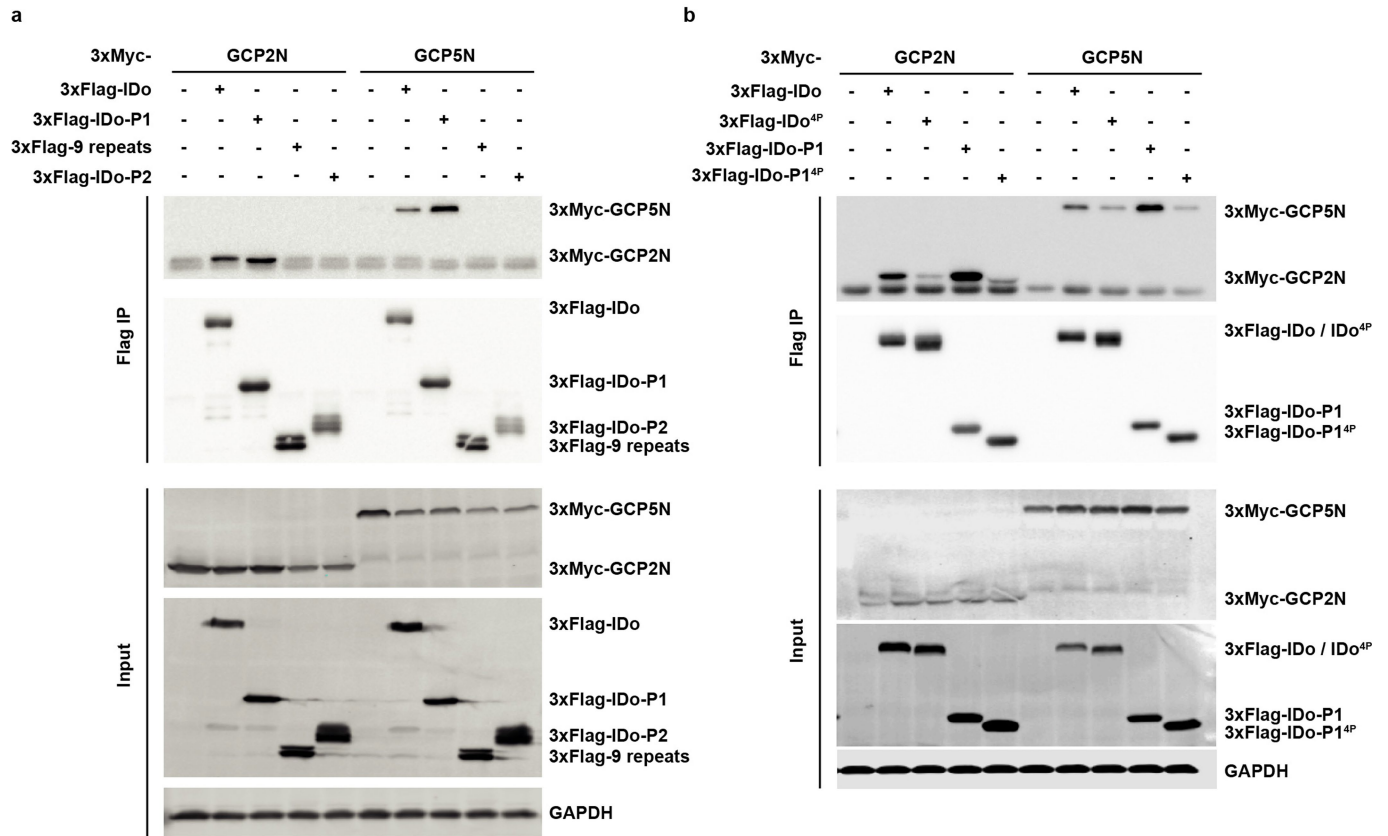
Extended Data Fig. 5 | Gallery of bulky amino acid side chains resolved in the γ -TuRC density. GCP-variant-specific bulky amino acid side chains are marked by an asterisk. Combinations of such side chains in close proximity are marked by two asterisks.



Extended Data Fig. 6 | See next page for caption.

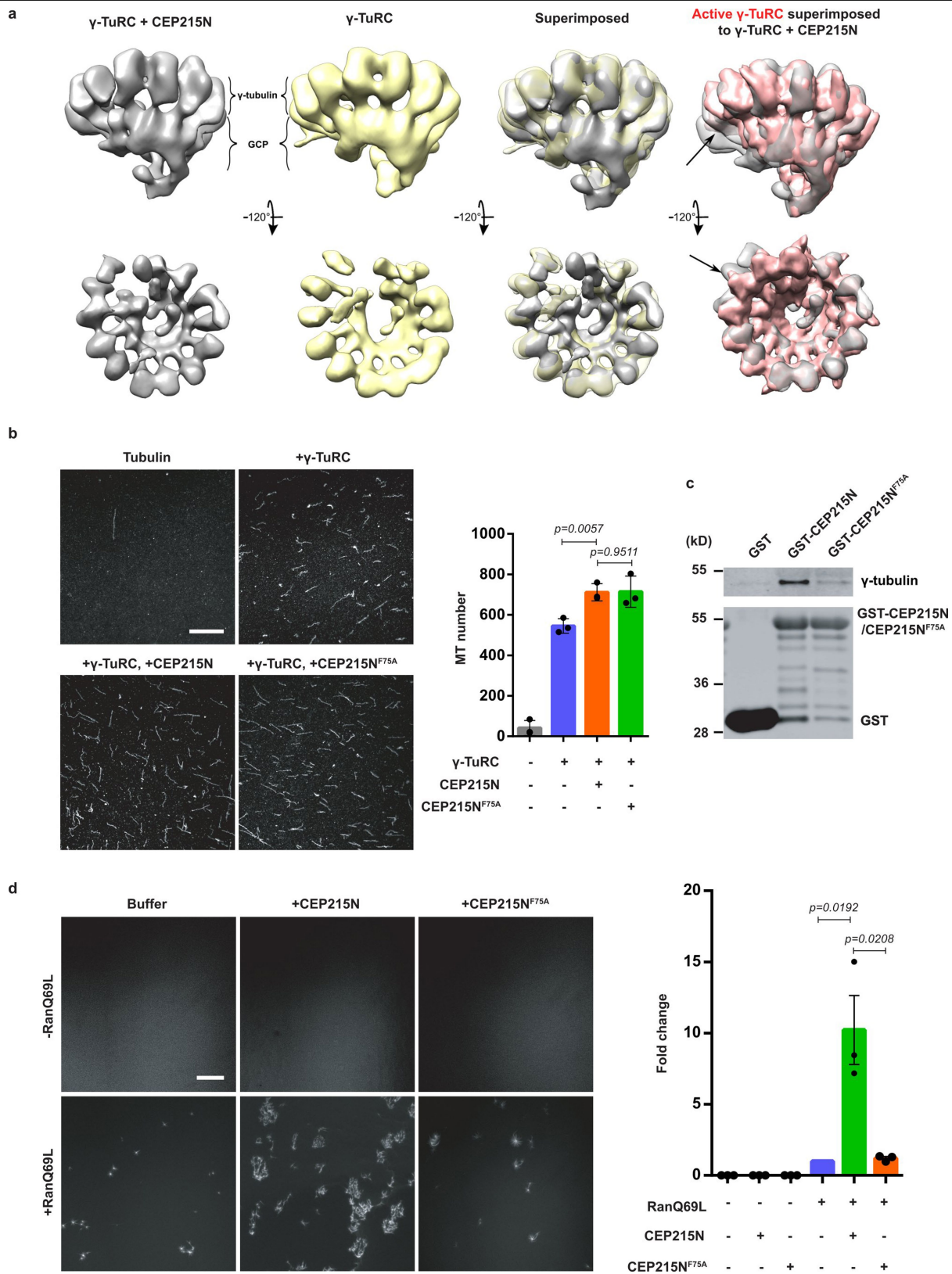
Extended Data Fig. 6 | Actin is a bona fide γ -TuRC component involved in MT nucleation. **a–d**, High-confidence homology models were docked with 10,000 randomly sampled starting positions and orientations into the envelope of the cryo-EM density. The number of fits is plotted against the respective correlation coefficients. **a**, γ -Tubulin and GCP4 served as positive controls. True positive fits correspond to redundant high-frequency fits with high correlation coefficients (dashed box). **b**, Ovalbumin served as negative control. No redundant high-frequency fits with high correlation coefficient indicative of a true positive match were observed. **c**, Unbiased fitting of NME7, NEDD1 and MOZART1 resulted in a distribution of correlation coefficients similar to the negative control, indicating no positive fit. **d**, Unbiased fitting of actin resulted in a distribution of correlation coefficients similar to the positive controls, clearly including a true positive fit (dashed box) shown in Fig. 1a. **e**, Purified γ -TuRCs from *X. laevis* egg extracts were resolved by SDS–PAGE and immunoblotted with anti- γ -tubulin, anti-GCP5 and anti-actin antibodies, which confirms that actin is associated with the purified γ -TuRC fraction. Representative blots are from three biologically independent experiments. For gel source data, see Supplementary Fig. 1. **f, g**, Indirect immunofluorescence of adsorbed γ -TuRC rings with antibodies directed against actin and GCP6 indicates colocalization of both proteins. Treatment with 1% SDS increased the antibody accessibility of actin located in the spatially confined interior of the

γ -TuRC. **f**, Representative fluorescence images together with magnified views from three biologically independent experiments. Scale bars, 20 nm. **g**, Percentage of colocalization events of actin and GCP6 normalized to the GCP6 signal. After treatment with 1% SDS, 40.2% of GCP6 signals colocalized with actin signals. In the absence of 1% SDS, colocalization events decreased to 6.4% owing to inaccessibility of the epitope. Data are mean \pm s.d. from $n = 3$ biologically independent experiments. *P* values were determined by unpaired two-sided *t*-test. **h**, Actin polymerization activity was tested for buffer with and without pyrene F-actin and 0.5 nM ARP2/3, 0.5 nM ARP2/3 with 15 nM VCA and 0.5 nM γ -TuRC with pyrene F-actin. The purified γ -TuRC is devoid of actin nucleation activity. The fluorescence intensity change was determined over time. Data are the mean of five independent experiments. **i**, γ -TuRC MT nucleation activity after pre-incubation with buffer, the actin-binding protein DNaseI and the preformed actin–DNaseI complex. $n = 3$ biologically independent experiments; data are mean \pm s.e.m. *P* values were determined by unpaired two-sided *t*-test. **j**, After 3 h incubation on ice, before the MT nucleation assay, samples were analysed by immunoblotting to confirm that equal amounts γ -TuRC were present in different experimental groups. Three biologically independent experiments were performed with similar results. For gel source data, see Supplementary Fig. 1.



Extended Data Fig. 7 | Co-immunoprecipitation of wild-type or mutant 3×Flag–GCP6 insertion domain fragments with the 3×Myc-tagged GCP2 and GCP5N termini. **a**, Co-immunoprecipitation of the GCP6 insertion domain (IDo; residues 606–1499), part 1 of the GCP6 insertion domain (IDo-P1; residues 606–1026), GCP6 insertion domain 9 repeats (IDo-9 repeats; residues 1027–1268) and part 2 of the GCP6 insertion domain (IDo-P2; residues 1269–1499) (as defined in Fig. 3a) with GCP2-N and GCP5-N. Immunoblotting was performed with indicated antibodies. Representative result is from three

independent experiments. For gel source data, see Supplementary Fig. 1. **b**, Co-immunoprecipitation of the GCP6 insertion domain or part 1 of the GCP6 insertion domain with residues V644, F706, F732 and Q783 mutated to proline (GCP6-IDo^{4P} and GCP6-IDo-P1^{4P}, respectively) with both GCP2-N and GCP5-N. Immunoblotting was performed with indicated antibodies. Three biologically independent experiments were performed with similar results. For gel source data, see Supplementary Fig. 1.



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | The CM1 motif of CEP215-N is not sufficient to activate the MT nucleation activity of γ -TuRC. **a**, γ -TuRC particles with or without a 600-fold excess of CEP215-N were analysed by negative-staining electron microscopy. The resulting 3D densities (grey, yellow) were superimposed and compared to a simulated density of the γ -TuRC in the extrapolated active conformation (red). Arrows indicate the most pronounced structural differences (transparent grey density not filled by red density) between the CEP215-N- γ -TuRC complex and the simulated density of the γ -TuRC in the extrapolated active conformation. **b**, γ -TuRC (0.5 nM) was incubated with an excess of CEP215-N or CEP215(F75A)-N (3 μ M). In vitro MT nucleation activity of the γ -TuRC incubated with buffer, glutathione *S*-transferase (GST), GST-CEP215-N or GST-CEP215(F75A)-N. $n = 3$ biologically independent experiments; data are mean \pm s.d. *P* values were determined by unpaired two-sided *t*-test.

Scale bar, 10 μ m. **c**, In vitro binding of γ -TuRC to purified and recombinant GST, CEP215-N and CEP215(F75A)-N. Immunoblots were probed with anti- γ -tubulin and anti-GST antibodies as shown. Three biologically independent experiments were performed with similar results. For gel source data, see Supplementary Fig. 1. **d**, MT nucleation activity in egg extracts induced by the addition of Ran(Q69L) and CEP215-N. The MT nucleation reaction was stopped after 15 min at 20 °C when Ran(Q69L) addition only induced a small number of MTs. Fold changes of total aster fluorescence intensity from ten random fields were quantified and normalized to the group adding Ran(Q69L) and without CEP215-N or CEP215(F75A)-N. $n = 3$ biologically independent experiments; data are mean \pm s.e.m. *P* values were determined by unpaired two-sided *t*-test. Scale bars, 50 μ m.

Extended Data Table 1 | Cryo-EM data collection, refinement and validation statistics

#1 γ -TuRC (EMDB-10491) (PDB 6TF9)				
Data collection and processing	dataset 1	dataset 2	dataset 3	dataset 4
Magnification	42,000	42,000	42,000	64,000
Voltage (kV)	300	300	300	300
Electron exposure (e-/Å ²)	35	42	51	57.4
Defocus range (μm)	-1.5 to -3.5	-1.5 to -3.5	-1.5 to -3.5	-1.0 to -3.0
Pixel size (Å)	2.1	2.1	2.1	1.35
Symmetry imposed	C1	C1	C1	C1
Initial particle images (no.)	774255	1174825	908652	2624596
Final particle images (no.)	17633	5621	7962	14880
Merged set of particles		46096		
Map resolution (Å)		4.8		
FSC threshold		0.143		
Map resolution range (Å)		4.5-6.0		
Refinement				
Initial model used (PDB code)	Homology models derived from GCP4 (3RIP), γ -tubulin (1Z5W) and Actin (1J6Z)			
Model resolution (Å)	4.75			
FSC threshold	0.143			
Model resolution range (Å)	4.75-6.89			
Map sharpening <i>B</i> factor (Å ²)	-180.0			
Model composition				
Non-hydrogen atoms	120125			
Protein residues	14974			
Ligands	0			
<i>B</i> factors (Å ²)				
Protein	141.25			
Ligand	—			
R.m.s. deviations				
Bond lengths (Å)	0.005			
Bond angles (°)	1.051			
Validation				
MolProbity score	1.99			
Clashscore	4.90			
Poor rotamers (%)	2.73			
Ramachandran plot				
Favored (%)	93.91			
Allowed (%)	5.85			
Disallowed (%)	0.24			

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | | |
|-------------------------------------|--|
| n/a | Confirmed |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of all covariates tested |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

- | | |
|-----------------|---|
| Data collection | Datasets for cryo-EM were collected with SerialEM (3.7). Data for CD analysis were collected using JASCO spectra manager (1.06.00). Data for immunoblots were collected either by LAS4000IR (2.1) or Image Studio (5.2). Datasets for negative stain EM were collected with SerialEM (1.0.0.1). Software WoRx (6.1.1) was used to collect the data for immunofluorescence experiment. Data for in vitro nucleation assay and egg extract experiment were collected by VisiView software. Data for Actin polymerization assay were collected by CLARIOstar. |
| Data analysis | EM data were processed using Relion 3.0-Beta, MotionCorr 2.0 and gCTF 1.06. All density map related figure were prepared in Chimera 1.13.1 and ChimeraX 0.9. Atomic modelling was done in Coot 0.8.9.2 and Coot 0.9. Molecular Flexible fitting was performed in QwikMD implemented in VMD 1.9.4.a35 and in NAMD 2.13. Refinement and final flexible fitting of the model was performed in website tool NAMDinator, and in Phenix 1.14. Data for γ -TuRC geometrical analysis of tubulins and for CD analysis were plotted in Gnuplot 5.2. Secondary structure predictions were done in website tool PSIPRED vs 3.0. Sequence alignment of GCP components was done in website tool PROMALS-3D. Analysis and vector visualisation of conformational change of γ -TuRC was performed in PyMol 2.2.3. Data for CD measurements were analyzed using JASCO spectra manager (1.06.00). MaxQuant software (1.6.2.6a) was used for LC-MS/MS analysis and database search. Fiji software (2.0.0-rc-46/1.50g) was used to analyze the microtubule nucleation and immunofluorescence data, and western blots images. Microsoft Office Excel (2011) was used to normalize the data before statistical analyses. GraphPad Prism 6 (6.01) was used for statistical analyses. Image Studio Lite (5.2.5) was used for western blot quantification in salt treatments. EMAN2 Project Manager (eman2.12) was used for negative stain EM particle picking, and the particles were classified and averaged using the IMAGIC-4D package. The GCP6 IDo domain alignment was performed with Clustal Omega build in Jalview software (2101-VM). |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Cryo-EM densities of the γ -TuRC filtered according to global or local resolution were deposited in the Electron Microscopy Data Bank (EMDB) under accession code EMD-10491. Atomic coordinates for the γ -TuRC were deposited at the Protein Data Bank (PDB) under accession code 6TF9. The original immunoblots and further source data from LFQ mass spectrometry (Fig. 2a), immunoblot quantification (Fig. 3c), geometric analysis of the atomic model (Figs. 3g, 4b, 4c, 4d), MT nucleation assays (Extended Data Figs. 1c, 6i, 9b, 9d), CD measurements (Extended Data Fig. 4e), unbiased structure-guided identification (Extended Data Fig. 6a-d), quantification of indirect immunofluorescence (Extended Data Fig. 6g) and Actin polymerization (Extended Data Fig. 6h) are included in the Supplementary Information. The raw cryo-EM micrograph movie stacks are available from the corresponding authors upon request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. Cryo-EM data were collected in a sufficient amount for reconstruction. For immunofluorescence and microtubule in vitro nucleation experiments, images were acquired in a number that is sufficient for statistic analysis (as specified in Methods). For negative stain EM, the images used for particle classification and average were enough to get clear views of the protein complexes. Sample size was chosen based on data variation.
Data exclusions	No data were excluded from the analyses.
Replication	All experiments, including the immunoprecipitation, mass spectrometry, and so on, were repeated at least three times to confirm the reproducibility. All replicates were successful.
Randomization	For immunofluorescence, mass spec analysis and microtubule in vitro nucleation experiments, images/data were acquired randomly. Other experiments were not related to randomization.
Blinding	Blinding was not applied to our study. For a number of experiments this was technically not possible: cyro-EM, IP experiment, CD measurments. For other experiments this was technically not feasible because of the large sample size: e.g. MT nucleation assays. In other cases, a finding was confirmed by an independent approach: e.g. actin in the γ -TuRC.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used

Anti- γ -tubulin rabbit polyclonal antibody, which was used for γ -TuRC purification, was homemade against the C-terminal peptide as described in the previous publication: doi: 10.1038/ncomms9722, and diluted 1:2000 in immunoblot in CEP215N pull-down assay. Anti- γ -tubulin mouse monoclonal antibody (GTU-88, T6557, Lot: 053K4839), which was used for immunoblotting, was

from Sigma-Aldrich and was diluted 1:4000. Rabbit anti-GCP2 polyclonal antibody (PA5-21433, Lot: QL2122721A) was from Thermo Fisher Scientific, and was diluted 1:1000 for immunoblotting. Rabbit anti-Xgrip109 (GCP3) polyclonal antibody, diluted 1:2000 in immunoblotting, was from Dr. Y. Zheng, and was previously described in the publication: DOI: 10.1083/jcb.141.3.675. Rabbit anti-GCP6 polyclonal antibody, diluted 1:2000 in immunoblotting, was from Dr. Y. Zheng, and was described in the previous publication: DOI: 10.1083/jcb.151.7.1525. Guinea pig anti-GCP6 polyclonal antibody for immunofluorescence was generated as described in in the previous publication: DOI: 10.1083/jcb.151.7.1525, and was diluted 1:200. Anti- β -Actin mouse monoclonal antibody (AC-74, A5316, Lot: 048M4843V) used in immunofluorescence (diluted 1:1000) and anti-Actin rabbit polyclonal antibody (A2066, Lot: 058M4812V) used in immunoblot (diluted 1:200) were from Sigma-Aldrich. Anti-GCP4 rabbit polyclonal antibody were raised against full-length purified GCP4, and was diluted 1:200 in immunoblotting. Anti-GCP5 mouse monoclonal antibody (E-1) was from Santa Cruz Biotechnology (sc-365837, Lot: E2014), and was diluted 1:500 in immunoblotting. Mouse anti-FLAG monoclonal antibody (9A3, 8146S, Lot: 3, diluted 1:500 in immunoblotting) and rabbit anti-GAPDH polyclonal antibody (14C10, 2118S, Lot: 10, diluted 1:1000 in immunoblotting) were from Cell Signaling Technology. Mouse monoclonal anti-c-Myc antibody (clone 9E10, M4439, Lot: 087M4765V, diluted 1:1000 in immunoblotting) was from Sigma-Aldrich. Secondary antibodies used in this study were: Donkey anti-Mouse Alexa Fluor 488-conjugated antibody (A21202, Lot: 1562298, diluted 1:500 in immunofluorescence) and goat anti-guinea pig Alexa Fluor® 555-conjugated antibody (A21435, Lot: 1711692, diluted 1:500 in immunofluorescence) are from Thermo Fisher Scientific; peroxidase-conjugated goat anti-mouse antibody (115-035-068, Jackson ImmunoResearch Laboratories, diluted 1:5000 in immunoblotting); donkey anti-mouse DyLight 680 (A10038, Lot: 1717043) and 800-conjugated antibodies (SA5-10172) are from Thermo Fisher Scientific, and both were diluted 1:5000 in immunoblotting; anti-rabbit DyLight 680-conjugated antibody (5366S, Lot: 3, Cell Signaling Technology, diluted 1:5000 in immunoblotting); IRDye 800CW Donkey anti-Rabbit IgG (926-32213, Lot: C61012-02, LI-COR Biosciences, diluted 1:5000 in immunoblotting).

Validation

Homemade rabbit anti- γ -tubulin rabbit polyclonal antibody and homemade rabbit anti-GCP4 rabbit polyclonal antibodies were validated previously in the publication: doi: 10.1038/ncomms9722

Homemade rabbit anti-GCP3 rabbit polyclonal antibody was validated previously in the publication: DOI: 10.1083/jcb.141.3.675

Homemade rabbit anti-GCP6 rabbit polyclonal antibody was validated previously in the publication: DOI: 10.1083/jcb.151.7.1525

Homemade Guinea pig anti-GCP6 polyclonal antibody was generated with the same epitope as homemade rabbit anti-GCP6 rabbit polyclonal antibody (DOI: 10.1083/jcb.151.7.1525), and was further validated by pilot experiments before using.

All commercial antibodies were validated by manufacturers for the species and applications, which can be found in the links below.

Anti- γ -tubulin mouse monoclonal antibody (GTU-88):
<https://www.sigmaaldrich.com/catalog/product/sigma/t6557?lang=de®ion=DE>

Rabbit anti-GCP2 polyclonal antibody: <https://www.thermofisher.com/antibody/product/GCP2-Antibody-Polyclonal/PA5-21433>

Anti- β -Actin mouse monoclonal antibody (AC-74):
<https://www.sigmaaldrich.com/catalog/product/sigma/a2228?lang=de®ion=DE>

Anti-Actin rabbit polyclonal antibody (A2066):
<https://www.sigmaaldrich.com/catalog/product/sigma/a2066?lang=de®ion=DE>

Anti-GCP5 mouse monoclonal antibody (E-1): <https://www.scbt.com/scbt/product/gcp5-antibody-e-1>

Mouse anti-FLAG monoclonal antibody (9A3):
<https://en.cellsignal.de/products/primary-antibodies/dykdddk-tag-9a3-mouse-mab-binds-to-same-epitope-as-sigma-s-anti-flag-m2-antibody/8146>

Rabbit anti-GAPDH polyclonal antibody (14C10):
<https://en.cellsignal.de/products/primary-antibodies/gapdh-14c10-rabbit-mab/2118>

Mouse monoclonal anti-c-Myc antibody (clone 9E10):
https://www.sigmaaldrich.com/catalog/product/sigma/m4439?lang=de®ion=DE&gclid=CjwKCAjwkqPrBRA3EiwAKdtwk-gHHmK21m-8cwwWs0O4fsjAXlMo3waZRhaAnZ8eq5lsEdpYzebNXBoC_qQQAvD_BwE

Donkey anti-Mouse Alexa Fluor 488-conjugated antibody: <https://www.thermofisher.com/antibody/product/Donkey-anti-Mouse-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21202>

Goat anti-guinea pig Alexa Fluor® 555-conjugated antibody: <https://www.thermofisher.com/antibody/product/Goat-anti-Guinea-Pig-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21435>

Peroxidase-conjugated goat anti-mouse antibody: <https://www.jacksonimmuno.com/catalog/products/115-035-068>

Donkey anti-mouse DyLight 680-conjugated antibody: <https://www.thermofisher.com/antibody/product/Donkey-anti-Mouse-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A10038>

Donkey anti-mouse DyLight 800-conjugated antibody: <https://www.thermofisher.com/antibody/product/Donkey-anti-Mouse-IgG-H-L-Cross-Adsorbed-Secondary-Antibody-Polyclonal/SA5-10172>

Anti-rabbit DyLight 680-conjugated antibody: <https://en.cellsignal.de/products/secondary-antibodies/anti-rabbit-igg-h-l-dylight-680-conjugate/5366>

IRDye 800CW Donkey anti-Rabbit IgG: <https://www.licor.com/bio/reagents/irdye-800cw-donkey-anti-rabbit-igg-secondary-antibody>

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

HEK293T cell line is as described (Panic, M, et. al. PLOS Genet. 11, e1005243 (2015).

Authentication

We verified the HEK293T cell line according to the morphology by light microscopy. Cell line was purchased from ATCC.

Mycoplasma contamination

HEK293T cell line was negative in the mycoplasma contamination test.

Commonly misidentified lines
(See [ICLAC](#) register)

No commonly misidentified lines were used in this study.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

Xenopus laevis (lab bread NASCO, USA), female, sexually mature, i.e. more than 9 months old.

Wild animals

No wild animals were used in this study.

Field-collected samples

No field-collected samples were used in this study.

Ethics oversight

The permission numbers for Xenopus laevis experiment are:35-9185.81/G-204/12 (Regierungspräsidium Karlsruhe, BW) ; 81-02.05.40.17.091 (LANUF Recklinghausen, NRW). Ethic justification of experiment is according to § 7a Abs. 2 Nr. 3 TierSchG.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

The structural basis for cohesin–CTCF-anchored loops

<https://doi.org/10.1038/s41586-019-1910-z>

Received: 21 February 2019

Accepted: 5 December 2019

Published online: 6 January 2020

Yan Li^{1,6}, Judith H. I. Haarhuis^{2,6}, Ángela Sedeño Cacciatore^{2,6}, Roel Oldenkamp², Marjon S. van Ruiten², Laureen Willems², Hans Teunissen³, Kyle W. Muir^{1,4*}, Elzo de Wit^{3*}, Benjamin D. Rowland^{2*} & Daniel Panne^{1,5*}

Cohesin catalyses the folding of the genome into loops that are anchored by CTCF¹. The molecular mechanism of how cohesin and CTCF structure the 3D genome has remained unclear. Here we show that a segment within the CTCF N terminus interacts with the SA2–SCC1 subunits of human cohesin. We report a crystal structure of SA2–SCC1 in complex with CTCF at a resolution of 2.7 Å, which reveals the molecular basis of the interaction. We demonstrate that this interaction is specifically required for CTCF-anchored loops and contributes to the positioning of cohesin at CTCF binding sites. A similar motif is present in a number of established and newly identified cohesin ligands, including the cohesin release factor WAPL^{2,3}. Our data suggest that CTCF enables the formation of chromatin loops by protecting cohesin against loop release. These results provide fundamental insights into the molecular mechanism that enables the dynamic regulation of chromatin folding by cohesin and CTCF.

The interphase genome is folded in 3D through the concerted action of cohesin and CTCF. These architectural factors regulate the interactions between regulatory elements along chromosomes to control gene expression^{1,4,5}. Cohesin is thought to catalyse genome folding through a process known as ‘loop extrusion’, which involves the formation of chromosome loops that are progressively enlarged^{6–10}. Genomic regions within which cohesin forms loops are also known as topologically associating domains (TADs), or loop domains. TADs are flanked by CTCF sites that are thought to act as barriers to the loop extrusion process^{11,12}. CTCF acts as such a boundary only when the 3′ ends of CTCF binding motifs are oriented towards the inside of the TAD^{9,13,14}. Consequently, only convergently oriented pairs of CTCF sites form CTCF-anchored loops^{15,16}.

This model is supported by genetic manipulation of cohesin and CTCF. Depletion of the core cohesin subunit SCC1 leads to loss of TADs^{12,17}. By contrast, depletion of the cohesin release factor WAPL increases the size of chromatin loops^{10,12,18}. CTCF depletion leads to a marked loss of CTCF-anchored loops^{11,12}. However, how CTCF can act as a directional boundary that controls cohesin loop extrusion remains unknown.

Here we have investigated the mechanism of cohesin interaction with CTCF, and how this interaction contributes to genome organization. We have identified an N-terminal segment of CTCF that directly engages the SA2–SCC1 subcomplex of cohesin. Our crystal structure of the SA2–SCC1–CTCF complex elucidates the molecular basis of the interaction. CTCF-anchored loops are abolished in mutants of key amino acids in the interface, but the accumulation of cohesin at CTCF binding sites across the genome is only partially impaired. In addition to its function as a translocation barrier, CTCF thus possesses a distinct loop-stabilizing activity, which is realized through a direct interaction

with cohesin. Furthermore, we observe intermolecular competition between CTCF and the cohesin release factor WAPL for this interface, which suggests a mechanism by which chromatin loop formation may be dynamically regulated.

Structure of the SA2–SCC1–CTCF complex

Previous data indicate that CTCF directly interacts with the SA2 subunit of the cohesin complex^{19,20}. To map this interaction, we produced a series of CTCF truncations as proteins fused to glutathione *S*-transferase (GST), and performed pulldown assays against a complex of SA2 and SCC1². CTCF fragments that contained amino acids 227–235 generally retained SA2–SCC1 on GST beads (Extended Data Fig. 1a, b). Isothermal calorimetry experiments further showed that the interaction is largely driven by amino acids 222–231 of CTCF, as the interaction involving this truncated CTCF retained an equilibrium dissociation constant ($K_d = 1.04 \pm 0.20 \mu\text{M}$) comparable to that of an extended CTCF construct ($K_d = 0.62 \pm 0.07 \mu\text{M}$) (Extended Data Fig. 1c, Extended Data Table 1a). To understand the molecular details, we produced crystals of the SA2–SCC1 complex in the presence of a peptide comprising the CTCF binding motif, and determined the structure by molecular replacement at a resolution of 2.7 Å (Extended Data Table 1b). An $F_o - F_c$ omit electron density Fourier map exhibited clear features that correspond to the CTCF peptide (Extended Data Fig. 1d).

The CTCF peptide is bound to the convex surface of SA2 (Fig. 1a, b). The CTCF binding surface is predominantly hydrophobic and composed of amino acids that are contributed by both SA2 and SCC1. The lead ‘anchoring’ amino acids of CTCF, which bury the largest solvent-accessible surface area upon binding, are Y226 and F228 (Fig. 1b). F228 inserts into a pocket comprising amino acids from SCC1 (S334, I337 and

¹European Molecular Biology Laboratory, Grenoble, France. ²Division of Gene Regulation, The Netherlands Cancer Institute, Amsterdam, The Netherlands. ³Division of Gene Regulation, Oncode Institute, The Netherlands Cancer Institute, Amsterdam, The Netherlands. ⁴MRC Laboratory of Molecular Biology, Cambridge, UK. ⁵Leicester Institute of Structural and Chemical Biology, Department of Molecular and Cell Biology, University of Leicester, Leicester, UK. ⁶These authors contributed equally: Yan Li, Judith H. I. Haarhuis, Ángela Sedeño Cacciatore.

*e-mail: kmuir@mrc-lmb.cam.ac.uk; e.d.wit@nki.nl; b.rowland@nki.nl; daniel.panne@le.ac.uk

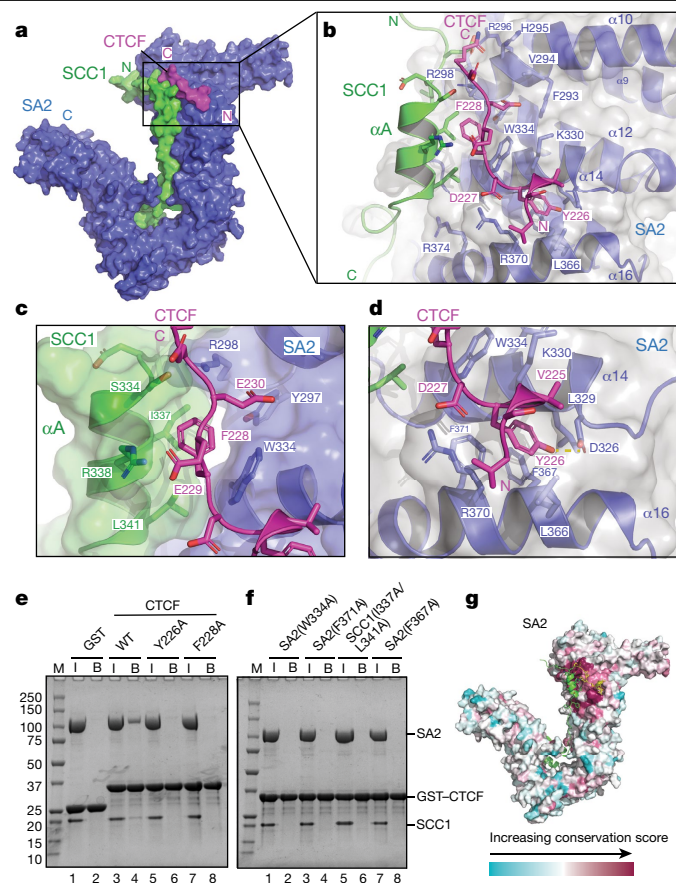


Fig. 1 | Structure of the SA2-SCC1-CTCF complex. **a**, Surface-rendered cartoon of the SA2-SCC1-CTCF complex, with components coloured in blue, green and magenta, respectively. C, C terminus; N, N terminus. **b**, Detailed view of the binding interface with SA2 residues in blue, SCC1 in green and CTCF in magenta. **c, d**, Details of the composite binding pocket around CTCF F228 (**c**) and CTCF Y226 (**d**). **e, f**, GST pull-down analysis of CTCF (**e**) and SA2 or SCC1 variants (**f**). B, bound fraction; I, input; M, molecular weight marker. Controls are shown in **e** (lanes 1 and 2). Experiments were done once. **g**, SA2 is surface-rendered and coloured according to sequence conservation.

L341) and SA2 (Y297 and W334) (Extended Data Fig. 1e). The hydroxyl group of Y226 hydrogen-bonds with D326 of SA2 in a deep hydrophobic pocket lined by L329, L366 and F367 (Fig. 1d). E229 and E230 of CTCF constitute secondary anchoring residues, which presumably contribute to binding specificity by forming salt bridges with R298 of SA2 and R338 of SCC1 (Fig. 1c). As CTCF engages a composite binding surface containing amino acids from SCC1 and SA2, previous mapping studies that used isolated SA2 may have been misleading²⁰.

Analysis of the CTCF binding interface

Mutagenesis of Y226A or F228A in CTCF abolished SA2-SCC1 binding in a GST pull-down assay (Fig. 1e). Likewise, the substitution of critical amino acid residues—including W334A, F371A or F367A in SA2 or I337A/L341A in SCC1—abolished CTCF binding (Fig. 1f). SA2 contains an 86-amino-acid motif termed the ‘stromalin conservative domain’^{21,22} or ‘conserved essential surface’ (CES)^{2,23}, which is conserved from fungi to mammals and coincides with the CTCF binding pocket. For simplicity, we refer to the composite SA2-SCC1 binding pocket as the CES. Mapping of sequence conservation onto the structure confirms that the CES is highly conserved (Fig. 1g, Extended Data Fig. 2a). A series of missense mutations are found in SA2 (also known as *STAG2*), SCC1 (also known as *RAD21*) and CTCF in various types of cancer²⁴. The mapping of mutation

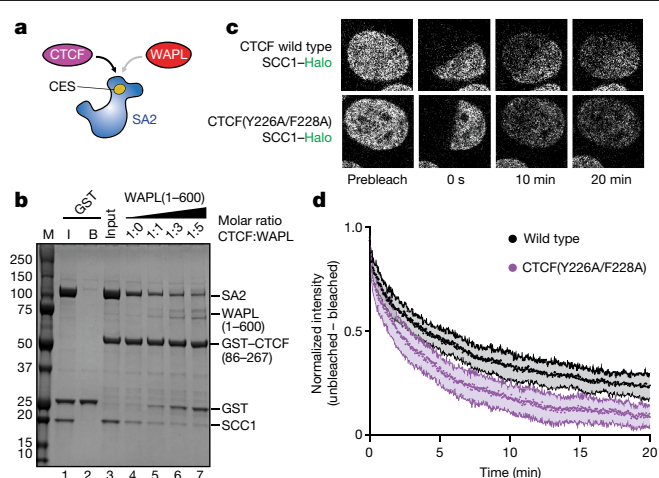


Fig. 2 | CTCF interaction stabilizes cohesin on DNA. **a**, Schematic of competition between CTCF and WAPL. **b**, Increasing amounts of WAPL residues 1–600 (WAPL(1–600)) (lane 4–7; molar ratios are indicated) were incubated with GST-CTCF and SA2-SCC1 and the bound fraction was analysed. Three independent experiments were done, with consistent results. A representative example is shown. **c**, Example images of cells used in **d** at the indicated time points after photobleaching. FRAP was performed in G1 cells (Extended Data Fig. 3d). **d**, Quantification of the FRAP experiments. Averages and standard deviations for 21 wild-type cells and 17 CTCF^{Y226A/F228A} cells, measured over 3 independent experiments.

frequencies onto the structure shows that amino acids that are largely buried in the interface are hotspots in cancer (Extended Data Fig. 2b).

Previous data indicate that the SA2-SCC1 complex interacts with multiple cohesin regulators^{2,23,25}. This includes two factors with opposing functions: WAPL, the general cohesin release factor, and shugoshin (SGO1), a factor that is crucial for the protection of centromeric cohesion during mitosis^{2,26–28}. This antagonism arises as a result of direct competition for binding to the CES of SA2-SCC1². As mutants reported to interfere with both SGO1 and WAPL binding cluster in the CES, we investigated whether these proteins bind to SA2-SCC1 by a mechanism comparable to that of CTCF. In SGO1, the reported CES-binding domain (amino acids 313–353) contains a conserved FGF-like motif that strongly resembles that of the CTCF peptide. Vertebrate WAPL also contains several FGF motifs in its N-terminal region that are potentially involved in cohesin regulation^{3,29}. A minimal fragment of WAPL capable of competing with SGO1 for access to the CES (amino acids 410–590) contains two such FGF motifs². We observed that a peptide that spans the second and third FGF motif of WAPL (amino acids 423–463) bound to SA2-SCC1 with a K_d of about 32.8 μ M (Extended Data Fig. 2c), whereas a peptide that comprises only the third motif bound more weakly (Extended Data Table 1a). The peptide containing the CES motif of CTCF therefore binds with higher affinity than do peptides that contain the WAPL motif(s).

CTCF stabilizes cohesin on chromatin

The observation that CTCF and WAPL can bind to the same surface on SA2-SCC1 raises the possibility that their interaction with the CES is mutually exclusive (Fig. 2a). To determine whether WAPL competes with CTCF for binding to the CES of SA2-SCC1, we performed GST-pull-down competition assays. Titration of WAPL residues 1–600 against a preformed complex of GST-CTCF and SA2-SCC1 depleted the latter from the beads (Fig. 2b). Similarly, titration of a peptide of SGO1 phosphorylated at T346—which has previously been reported to preclude WAPL binding²—also displaced SA2-SCC1 from GST-CTCF (Extended Data Fig. 2d). Hence, the CES of SA2-SCC1 is a general interaction hub for multiple regulators of cohesin (Extended Data Fig. 2e). Whereas

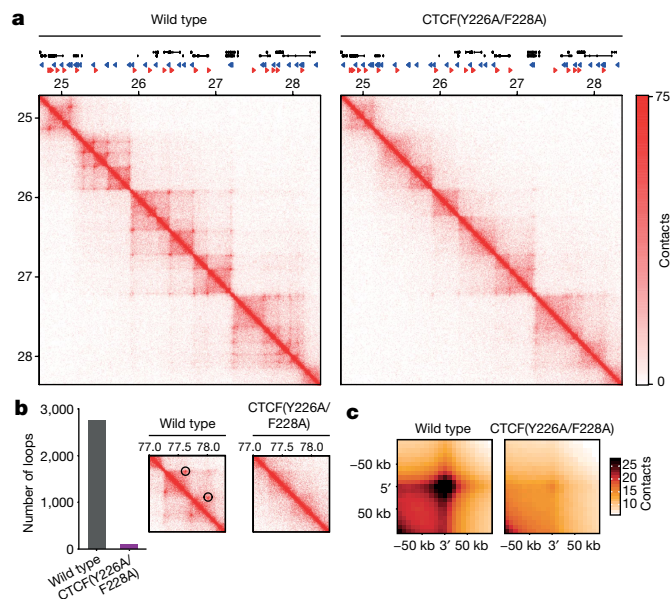


Fig. 3 | CTCF–CES interaction is required for CTCF-anchored loops. **a**, Hi-C contact matrices of the *HOXA* locus at 10-kb resolution, normalized to 100 million contacts per sample. Genes and CTCF sites are depicted above the contact matrices. **b**, Genome-wide quantification of loops using HICCUPS¹⁵. The inset shows an example of called loops for a region of chromosome 16. **c**, Aggregate peak analysis for the loops defined genome-wide in wild-type cells. The Hi-C signal is averaged across these locations for both cell lines.

SGO1 precludes WAPL binding (thus stabilizing centromeric cohesin in mitosis), CTCF could exert a similar function at CTCF sites in interphase.

To test whether the CTCF–CES interaction stabilizes cohesin on chromatin, we mutated the endogenous allele of *CTCF* in the human haploid HAP1 cell line using CRISPR–Cas9 technology. We thereby obtained HAP1 cells that contained the *CTCF*^{Y226A/F228A} mutation as their sole copy of *CTCF* (Extended Data Fig. 3a, b). These cells displayed no obvious proliferation defects. To study the consequences of the *CTCF* mutations on cohesin turnover on chromatin, we endogenously tagged the core cohesin subunit SCC1 with a Halo tag in both wild-type and *CTCF*^{Y226A/F228A} cells (Extended Data Fig. 3c), and performed fluorescence recovery after photobleaching (FRAP) experiments. In wild-type cells, we found that—over a period of 20 min—a fraction of the fluorescent cohesin population did not recover. However, in *CTCF*^{Y226A/F228A} cells we observed a near-complete recovery by FRAP, which demonstrates that cohesin is more mobile in these cells (Fig. 2c, d). The CTCF–CES interaction therefore stabilizes a subpopulation of cohesin on chromatin.

Loops require CTCF–CES binding

To investigate the role of the cohesin–CTCF interaction in chromosome organization, we generated chromosome conformation capture (Hi-C) profiles of wild-type and *CTCF*^{Y226A/F228A} cells. Wild-type HAP1 cells displayed clear loops connecting CTCF sites (Fig. 3a, Extended Data Fig. 4). However, Hi-C matrices of *CTCF*^{Y226A/F228A} cells revealed a robust ablation of CTCF-anchored loops (Fig. 3a). By systematically scoring the number of loops, we found that in the *CTCF*^{Y226A/F228A} mutant the vast majority of detectable loops across the genome were lost (from 2,756 in the wild-type cells, to 98 in the mutant cells) (Fig. 3b). An aggregate peak analysis, which quantifies the contact frequency of all the loops identified in wild-type cells, likewise showed a marked loss of these contacts (Fig. 3c, Extended Data Fig. 4d).

CTCF sites not only lie at the bases of CTCF-anchored loops, but also form the boundaries of TADs (Extended Data Fig. 4a). We then assessed the effect of the *CTCF*^{Y226A/F228A} mutation on TADs, and found that these

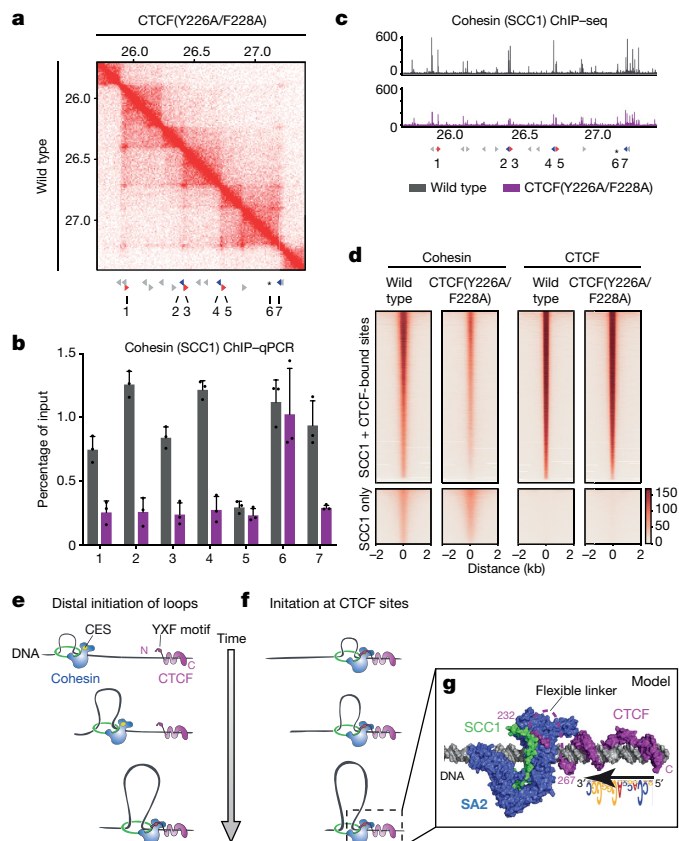


Fig. 4 | CTCF–CES interaction promotes localization of cohesin to CTCF sites.

a, Hi-C contact matrix of a region of chromosome 7 at 10-kb resolution. CTCF sites are depicted below; those selected for qPCR are shown in colour (forward motifs in red, reverse motifs in blue). The numbers underneath indicate the loci used for qPCRs in **b**. Locus 6 (indicated with *) is the *HOTAIRM1* transcription start site. **b**, ChIP–qPCR analysis of SCC1 (cohesin) at the loci depicted in **a**. The mean of three independent ChIP experiments is shown with standard deviations. **c**, ChIP–seq tracks for SCC1 of the same region of chromosome 7 as is depicted by Hi-C in **a**. The ChIP–qPCR loci of **b** are depicted below. **d**, ChIP–seq heat map of the cohesin subunit SCC1 (left) and CTCF (right). The depicted sites are selected for being bound in wild-type cells by both SCC1 and CTCF (top), or only by SCC1 (bottom). **e**, Cohesin-mediated looping initiates at distal sites until encounter of the N-terminal end of CTCF. **f**, Cohesin-mediated looping starts at CTCF sites. **g**, Molecular model of CTCF and SA2–SCC1 bound to DNA (grey). The YXF motif is separated by a flexible linker spanning residues 232–267 (magenta dotted line) to the C-terminal DNA-binding domain of CTCF.

structures were—to a considerable degree—still present in *CTCF*^{Y226A/F228A} cells but that they have less-clear edges (Fig. 3a). Aggregate TAD analysis further confirmed that TAD-like structures do exist in *CTCF*^{Y226A/F228A} cells, but that these structures have less-clear boundaries (Extended Data Fig. 4b, c, e, f) and completely lack CTCF loops at their corners (Extended Data Fig. 4b, e). Our results therefore support the notion that in *CTCF*^{Y226A/F228A} cells cohesin can form the loops along DNA that make up the contacts within TADs, but that cohesin is not stabilized at CTCF sites to allow for the formation or maintenance of CTCF-anchored loops.

Cohesin localization to CTCF sites

To assess whether the CTCF–CES interaction affects cohesin abundance at loop anchors, we performed chromatin immunoprecipitation with quantitative PCR (ChIP–qPCR) experiments. We selected CTCF sites at the base of loops (Fig. 4a, Extended Data Fig. 5a) and found that in the *CTCF*^{Y226A/F228A} mutant the abundance of cohesin was reduced at the majority of these loci. By contrast, cohesin levels at a nearby locus

that did not contain a CTCF site were not affected (Fig. 4b, Extended Data Fig. 5b). CTCF binding to the corresponding CTCF sites was also largely unaffected (Extended Data Fig. 5c–e). We then assessed cohesin distribution genome-wide by chromatin immunoprecipitation with sequencing (ChIP-seq) and found that the *CTCF*^{Y226A/F228A} mutation decreased cohesin localization to CTCF sites, but had little-to-no effect on cohesin localization at unrelated sites (Fig. 4c, d). Although cohesin levels at CTCF sites were reduced in *CTCF*^{Y226A/F228A} cells, cohesin was—to a considerable degree—still present at CTCF sites. Our data therefore support a model in which CTCF influences cohesin in two ways: (i) it halts cohesin at CTCF sites and (ii) it stabilizes cohesin at the base of CTCF-anchored loops. The former function could be important for defining TAD boundaries. The binding of CTCF to the CES of cohesin could affect the latter function and may thereby prevent the disruption of CTCF-anchored loops.

To evaluate the consequences of the loss of CTCF-anchored loops on gene expression, we performed RNA-sequencing analyses. The *CTCF*^{Y226A/F228A} mutation affected the expression of more than 2,000 genes. Although the number of genes that were upregulated was comparable to the number of genes that were downregulated, the most strongly affected genes were more frequently downregulated (Extended Data Fig. 7a). Thus, the interface of CTCF formed by Y226 and F228 and (by extension) cohesin–CTCF anchored loops are apparently key to correct expression of these genes. Despite this effect on gene expression and the loss of virtually all CTCF-anchored loops, cells that contain only this mutant form of CTCF are viable. CTCF has previously been shown to be essential for viability of mouse embryonic stem cells¹¹. We therefore tested whether CTCF is essential for the viability of HAP1 cells, and found that CTCF depletion mediated by short interfering RNA was lethal to both control HAP1 cells and *CTCF*^{Y226A/F228A} mutant cells (Extended Data Fig. 7b, c). Thus, CTCF has essential roles that are apparently independent of CES engagement and the formation of CTCF-anchored loops in these cells.

Identification of CES ligands

To investigate the prevalence of the CES-binding factors, we compiled an alignment of known cohesin partners and derived a regular expression motif (Extended Data Fig. 2e, f). We used this motif to query the human and budding yeast proteomes for proteins that contain similar binding motifs³⁰. From the set of nuclear proteins that arose from this search, we were able to identify known cohesin regulators as well as several additional potential binding factors. We generated peptide arrays that bear these sequences and assayed the binding of SA2–SCC1, using an SA2(F371A)–SCC1 mutant complex as a negative control. We observed clear signal for the CTCF peptide that spans amino acids 222–231, which was abolished in the SA2(F371A)–SCC1 mutant (Extended Data Fig. 7d, e). A CTCF(Y226F) mutant showed approximately 1.5-fold reduced binding, apparently due to loss of the hydrogen bond between the hydroxyl group of CTCF Y226 and D326 of SA2 (Extended Data Table 2). Consistent with our pull-downs, the CTCF(Y226A), CTCF(F228A) and CTCF(Y226A/F228A) peptide variants did not retain SA2–SCC1. The WAPL peptides showed considerably weaker binding as compared to CTCF, and we could not detect binding for ligands such as SGO1 (Extended Data Table 1a, Methods). Robust binding was observed for MCM3 (a subunit of the replicative helicase), SYCP3 (a component of the synaptonemal complex), ZGPAT (a transcriptional repressor) and CENPU (a subunit of the inner kinetochore). Thus, the CES of SA2–SCC1 potentially facilitates cohesin regulation for a number of functionally divergent chromosomal processes.

Discussion

Our study reveals that CTCF binds to a CES on the SA2–SCC1 subcomplex of cohesin. The ablation of this interaction results in a near-complete

loss of CTCF–cohesin anchored loops. Thus, CTCF does not simply present a passive barrier to cohesin-mediated loop extrusion, but specifically interacts with the CES to stabilize cohesin at these loci and to prevent loop disruption. Accordingly, impairment of the CTCF–CES interaction renders cohesin more dynamic (Fig. 2c, d).

SA2 and SCC1, as well as *CTCF*, are frequently mutated in a number of tumour types³¹ and the mutations cluster in the CES (Extended Data Fig. 2b). Therefore, the dysregulation of chromatin looping may be causally related to carcinogenesis^{32,33}.

We envisage two possible scenarios for the formation of CTCF-anchored chromatin loops. In the first model (Fig. 4e), cohesin initiates loop enlargement at distal chromatin loci. These cohesin complexes remain dynamic because the cohesin release factor WAPL directly binds to cohesin by engaging the CES^{2,3,29} and PDS5^{12,27,34}, and promotes the opening of cohesin rings at the SMC3–SCC1 interface^{35–37}. Alternatively, loop enlargement commences at CTCF sites³⁸. Cohesin then catalyses DNA looping at these sites because CTCF counteracts DNA release (Fig. 4f). These models are not necessarily mutually exclusive, as a cohesin complex that initiates looping in the former mode may well be converted into the latter upon encountering CTCF. As CTCF directly competes with WAPL for binding to the CES (Fig. 2a, b), we suggest that this interaction stabilizes chromatin loops.

We propose a model for how cohesin and CTCF co-associate on DNA (Fig. 4g). Our model indicates that cohesin engages CTCF only when approaching the N terminus of CTCF. Specifically, the 34-amino-acid flexible linker that connects the YXF motif to the first DNA-binding zinc finger of human CTCF is sufficiently long to allow SA2–SCC1 DNA binding towards the N, but not the C, terminus of CTCF (Fig. 4g), thus confirming previous mapping studies³⁹. Stabilization of cohesin by engagement of the CTCF N terminus may explain why TAD boundaries arise preferentially when CTCF binding sites are convergently oriented^{9,13–16,39,40}. If an individual cohesin complex anchors itself at the N terminus of CTCF, and then reels in DNA until it encounters a cohesin that is likewise reeling from the opposite CTCF site, this would bring together CTCF sites³⁸. Loop formation by the related *Saccharomyces cerevisiae* condensin complex appears to involve a DNA anchoring function of its HEAT-repeat subunit Ycg1, a paralogue of human SA2 and *Saccharomyces cerevisiae* Scc3^{41–43}. These different complexes may therefore use a similar anchoring principle to build loops and provide structure to genomes. As the CES interface is conserved between isoforms of SA, we anticipate that ligand binding will affect all cohesin variants in a similar manner. Similarly, this interface is also conserved through Scc3 in fungi, despite the absence of CTCF in these organisms. The CES therefore is likely to represent an ancient interaction hub on cohesin.

The observation that CTCF–CES interaction controls DNA looping indicates that this aspect of cohesin function can be regulated by an F/YXF motif containing cohesin ligands. A number of other genome regulatory factors contain F/YXF motifs—including SGO1 (Extended Data Fig. 7d, e), which protects centromeric chromatid cohesion by antagonizing WAPL binding to the CES of SA2–SCC1². We therefore predict that a number of proteins that contain F/YXF motifs engage the CES and thereby modulate the ability of cohesin to catalyse genome folding in functionally divergent chromosomal processes.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1910-z>.

1. Dekker, J. & Mirny, L. The 3D genome as moderator of chromosomal communication. *Cell* **164**, 1110–1121 (2016).

2. Hara, K. et al. Structure of cohesin subcomplex pinpoints direct shugoshin–Wapl antagonism in centromeric cohesion. *Nat. Struct. Mol. Biol.* **21**, 864–870 (2014).
3. Shintomi, K. & Hirano, T. Releasing cohesin from chromosome arms in early mitosis: opposing actions of Wapl–Pds5 and Sgo1. *Genes Dev.* **23**, 2224–2236 (2009).
4. Merkenschlager, M. & Nora, E. P. CTCF and cohesin in genome folding and transcriptional gene regulation. *Annu. Rev. Genomics Hum. Genet.* **17**, 17–43 (2016).
5. Rowley, M. J. & Corces, V. G. Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* **19**, 789–800 (2018).
6. Yatskevich, S., Rhodes, J. & Nasmyth, K. Organization of chromosomal DNA by SMC complexes. *Annu. Rev. Genet.* **53**, 445–482 (2019).
7. Alipour, E. & Marko, J. F. Self-organization of domain structures by DNA-loop-extruding enzymes. *Nucleic Acids Res.* **40**, 11202–11212 (2012).
8. Fudenberg, G. et al. Formation of chromosomal domains by loop extrusion. *Cell Rep.* **15**, 2038–2049 (2016).
9. Sanborn, A. L. et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl Acad. Sci. USA* **112**, E6456–E6465 (2015).
10. Haarhuis, J. H. I. et al. The cohesin release factor WAPL restricts chromatin loop extension. *Cell* **169**, 693–707 (2017).
11. Nora, E. P. et al. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* **169**, 930–944 (2017).
12. Wutz, G. et al. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J.* **36**, 3573–3599 (2017).
13. Guo, Y. et al. CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* **162**, 900–910 (2015).
14. de Wit, E. et al. CTCF binding polarity determines chromatin looping. *Mol. Cell* **60**, 676–684 (2015).
15. Rao, S. S. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).
16. Vietri Rudan, M. et al. Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep.* **10**, 1297–1309 (2015).
17. Rao, S. S. P. et al. Cohesin loss eliminates all loop domains. *Cell* **171**, 305–320 (2017).
18. Gassler, J. et al. A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *EMBO J.* **36**, 3600–3618 (2017).
19. Rubio, E. D. et al. CTCF physically links cohesin to chromatin. *Proc. Natl Acad. Sci. USA* **105**, 8309–8314 (2008).
20. Xiao, T., Wallace, J. & Felsenfeld, G. Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity. *Mol. Cell. Biol.* **31**, 2174–2183 (2011).
21. Pezzi, N. et al. STAG3, a novel gene encoding a protein involved in meiotic chromosome pairing and location of STAG3-related genes flanking the Williams-Beuren syndrome deletion. *FASEB J.* **14**, 581–592 (2000).
22. Orgil, O. et al. A conserved domain in the Scc3 subunit of cohesin mediates the interaction with both Mod1 and the cohesin loader complex. *PLoS Genet.* **11**, e1005036 (2015).
23. Roig, M. B. et al. Structure and function of cohesin's Scc3/SA regulatory subunit. *FEBS Lett.* **588**, 3692–3702 (2014).
24. Forbes, S. A. et al. COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Res.* **45**, D777–D783 (2017).
25. Beckouët, F. et al. Releasing activity disengages cohesin's Smc3/Scc1 interface in a process blocked by acetylation. *Mol. Cell* **61**, 563–574 (2016).
26. Gandhi, R., Gillespie, P. J. & Hirano, T. Human Wapl is a cohesin-binding protein that promotes sister-chromatid resolution in mitotic prophase. *Curr. Biol.* **16**, 2406–2417 (2006).
27. Kueng, S. et al. Wapl controls the dynamic association of cohesin with chromatin. *Cell* **127**, 955–967 (2006).
28. Liu, H., Rankin, S. & Yu, H. Phosphorylation-enabled binding of SGO1–PP2A to cohesin protects sororin and centromeric cohesion during mitosis. *Nat. Cell Biol.* **15**, 40–49 (2013).
29. Ouyang, Z. et al. Structure of the human cohesin inhibitor Wapl. *Proc. Natl Acad. Sci. USA* **110**, 11355–11360 (2013).
30. Krystkowiak, I. & Davey, N. E. SLIMSearch: a framework for proteome-wide discovery and annotation of functional modules in intrinsically disordered regions. *Nucleic Acids Res.* **45**, W464–W469 (2017).
31. Lawrence, M. S. et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505**, 495–501 (2014).
32. Flavahan, W. A. et al. Insulator dysfunction and oncogene activation in *IDH* mutant gliomas. *Nature* **529**, 110–114 (2016).
33. Hnisz, D. et al. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* **351**, 1454–1458 (2016).
34. Ouyang, Z., Zheng, G., Tomchick, D. R., Luo, X. & Yu, H. Structural basis and IP6 requirement for Pds5-dependent cohesin dynamics. *Mol. Cell* **62**, 248–259 (2016).
35. Chan, K. L. et al. Cohesin's DNA exit gate is distinct from its entrance gate and is regulated by acetylation. *Cell* **150**, 961–974 (2012).
36. Buheitel, J. & Stemmann, O. Prophase pathway-dependent removal of cohesin from human chromosomes requires opening of the Smc3–Scc1 gate. *EMBO J.* **32**, 666–676 (2013).
37. Eichinger, C. S., Kurze, A., Oliveira, R. A. & Nasmyth, K. Disengaging the Smc3/kleisin interface releases cohesin from *Drosophila* chromosomes during interphase and mitosis. *EMBO J.* **32**, 656–665 (2013).
38. Sedeño Cacciatore, Á. & Rowland, B. D. Loop formation by SMC complexes: turning heads, bending elbows, and fixed anchors. *Curr. Opin. Genet. Dev.* **55**, 11–18 (2019).
39. Tang, Z. et al. CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell* **163**, 1611–1627 (2015).
40. Nagy, G. et al. Motif oriented high-resolution analysis of ChIP-seq data reveals the topological order of CTCF and cohesin proteins on DNA. *BMC Genomics* **17**, 637 (2016).
41. Kschonsak, M. et al. Structural basis for a safety-belt mechanism that anchors condensin to chromosomes. *Cell* **171**, 588–600 (2017).
42. Ganji, M. et al. Real-time imaging of DNA loop extrusion by condensin. *Science* **360**, 102–105 (2018).
43. Li, Y. et al. Structural basis for Scc3-dependent cohesin recruitment to chromatin. *eLife* **7**, e38356 (2018).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

Constructs, protein expression and purification

The human SA2 fragment amino acids 80–1060, cloned into pGEX-6P and codon-optimized for expression in *Escherichia coli*, was obtained from H. Yu. The construct encodes an N-terminal GST tag and C-terminal SA2 separated by a PreScission protease cleavage site. A plasmid encoding SCC1 was obtained from J.-M. Peters. SA2 was co-expressed with an N-terminally 6×His-tagged fragment of SCC1 spanning residues 281–420 cloned into the NcoI–NotI sites of a pACYC-Duet-1 vector (Merck Millipore). CTCF constructs were cloned into the BamHI and NotI sites of pGEX-6P1. Mutagenesis was performed using a QuikChange Lightning site-directed mutagenesis kit (Agilent). All the proteins were expressed in *E. coli* BL21(DE3) by autoinduction⁴⁴. Cells were grown at 37 °C until an optical density at 600 nm ($OD_{600\text{nm}}$) = 0.6 and then shifted to 18 °C for 16 h. Cells were collected with a JLA-8.1 rotor (Beckman) and washed once with ice-cold PBS buffer. Pellets were resuspended in buffer 1 (40 mM TRIS, pH 7.5, 500 mM NaCl and 0.5 mM TCEP), lysed using a microfluidizer (Microfluidics) and centrifuged at 4 °C for 1 h at 15,000 rpm using JA-20/14 rotors (Beckman).

The GST- and His-tagged SA2–SCC1 complex was applied to Co^{2+} conjugated IMAC sepharose resin (GE Healthcare) using a Minipuls3 peristaltic pump (Gilson), washed with buffer 1 supplemented with 20 mM imidazole and eluted using buffer 1 supplemented with 300 mM imidazole. Co^{2+} eluate was then bound to Glutathione Sepharose 4 Fast Flow resin (GE Healthcare) using a Minipuls3 peristaltic pump (Gilson), washed with buffer 1 and eluted by adding 10 mM reduced L-Glutathione (Sigma-Aldrich) into buffer 1. The GST tag was cleaved by PreScission protease (EMBL core facilities) during overnight incubation at 4 °C. Cleaved protein was concentrated using an Amicon Ultra-15 concentrator (Millipore) and applied to a MonoQ 5/50 GL column (GE Healthcare) in buffer 2 (40 mM TRIS, pH7.5, 150 mM NaCl and 0.5 mM TCEP) and eluted via a linear gradient of buffer 2 containing 1 M NaCl and further purification using a HiLoad 16/60 Superdex 200 prep-grade column (GE Healthcare) in buffer 3 (20 mM TRIS, pH7.7, 300 mM NaCl and 5 mM TCEP). The final purified proteins were concentrated using an Amicon Ultra-15 concentrator (Millipore) and flash-frozen in liquid N_2 for storage at –80 °C.

Crystallization and structure determination

Crystals of SA2(80–1060) in complex with SCC1 amino acids 281–420 (otherwise denoted the SA2–SCC1 complex) were grown by hanging-drop vapour diffusion at 20 °C by mixing equal volumes of protein at 8 mg ml^{–1} and crystallization solution containing 0.06 M Morpheus Divalents mix, 0.1 M Morpheus buffer system 2, 48% (v/v) Morpheus EOD_P8K (Molecular Dimensions). Crystals were soaked for 24–48 h with a peptide (obtained from peptid.de) including amino acid residues 222–231 of CTCF (Uniprot ID Q8NI51; DVSVDYDFEE). Crystals were cryo-protected by adding 15% glycerol to the well solution and flash-frozen in liquid nitrogen.

Diffraction data for all crystals were collected at 100 K at an X-ray wavelength of 0.966 Å at beamline ID30A-1/MASSIF-1⁴⁵ of the European Synchrotron Radiation Facility, with a Pilatus3 2M detector, using automatic protocols for the location and optimal centring of crystals⁴⁶. The beam diameter was selected automatically to match the crystal volume of highest homogeneous quality⁴⁷. Data were processed with XDS⁴⁸ and imported into CCP4 format using AIMLESS⁴⁹.

The structure was determined by molecular replacement using Phaser⁵⁰. A final model was produced by iterative rounds of manual model-building in Coot⁵¹ and refinement using PHENIX⁵². The

CTCF-containing model was refined to a resolution of 2.7 Å with an R_{work} and an R_{free} of 25% and 27%, respectively (Extended Data Table 1b). Analysis by MolProbity⁵³ showed that there are no residues in disallowed regions of the Ramachandran plot and the all-atom clash score was 7.2. The model shown in Fig. 4f was generated by superposition on DNA of SA2–SCC1–CTCF (RCSB Protein Data Bank code (PDB) 6QNX) with DNA-bound *Saccharomyces cerevisiae* SCC3–SCC1 (PDB 6H8Q)⁴³ and a composite model of DNA-bound CTCF zinc fingers assembled from PDB 5YEF and PDB 5YEL⁵⁴.

GST pulldowns and peptide arrays

For GST pulldowns, 10 μM GST-tagged CTCF constructs and 2.5 μM SA2–SCC1 were mixed in 50 μl buffer 4 (20 mM TRIS, pH7.7, 300 mM NaCl and 0.5 mM TCEP) + 0.1% Tween-20 containing 25 μl of a 50% slurry of GST sepharose beads per reaction. For WAPL and SGO1 competition assays, 2.5 μM GST-tagged CTCF (86–267) was incubated with 1 μM of SA2–SCC1 and increasing concentrations of WAPL (1–600) or a SGO1 phosphorylated at T346 peptide spanning amino acids 331–349 (molar ratios are indicated in each figure), under reaction conditions that were otherwise identical to GST pulldowns. Reactions were incubated at for 1 h at 4 °C. Twenty-five microlitres of the reaction were withdrawn as the reaction input and the remainder was washed 5 times with 500 μl of buffer 4 + 0.1% Tween-20. Samples were boiled in 1× SDS sample loading buffer (NEB) for 5 min to obtain the bound fraction, followed by SDS–PAGE analysis.

Isothermal calorimetry (ITC) was performed using a MicroCal iTC 200 (Malvern Panalytical) at 25 °C. SA2–SCC1 and the CTCF, SGO1 and WAPL peptide ligands were dialysed overnight at 4 °C against 20 mM TRIS, pH7.7, 150 mM NaCl, 0.5 mM TCEP. For each titration, 300 μl of 50 μM SA2–SCC1 was added to the calorimeter cell. The concentration of peptides was adjusted to 500 μM and injected into the sample cell as 16× 2.5-μl syringe fractions. Results were analysed and displayed using the Origin 7.0 software package supplied with the instrument. Data were analysed using the one-site binding model.

Peptide arrays, with an area of 3 cm², were obtained from R. Volkmer (<http://immunologie.charite.de>). Arrays were washed with 100% ethanol for 5 min on a shaker at 21 °C, followed by 3 washes, for a total of 10 min in TBS-T buffer (50 mM Tris pH7.5, 150 mM NaCl and 0.05% Tween-20). For the blocking step, arrays were incubated in 1× blocking buffer (Sigma B6429) for 3 h at 21 °C, followed by 3 washes in TBS-T for a total of 10 min. SA2–SCC1 and SA2(F371A)–SCC1 were added to 1× blocking buffer at a final concentration of 1.2 μM and incubated with the array overnight at 4 °C under gentle agitation. The membrane was washed 3 times (1× 30 s, and then 2× 5 min) at 21 °C. The anti 6× anti-poly His–HRP antibody (Sigma A-7058) was diluted 1:2,000 in 1× blocking buffer and incubated with the arrays for 1 h at 21 °C. The array was washed 3 times (1× 30 s, and then 2× 5 min) and developed by addition of 3,3′-diaminobenzidine (Sigma D4293) for 1 min followed by quenching in deionized H₂O. To measure non-specific binding of the anti-6×His antibody, all steps were identical except that no SA2–SCC1 protein solution was added to 1× blocking buffer during the overnight-incubation step. Arrays were imaged with a BioRad Gel Doc XR+ documentation system. Spot intensities were measured using ImageJ 1.52k. Three independent experiments were done and the apparent dissociation constants determined by normalization with ITC data from CTCF(222–231) (Extended Data Table 1a).

Genome editing and cell culture

Cells were cultured in Iscove's modified Dulbecco's medium supplemented with 10% FCS (Clontech), 1% penicillin–streptomycin (Invitrogen) and 1% UltraGlutamin (Lonza). The guide (g)RNA targeting exon 1 of *CTCF* was designed and annealed into pX330 (primer, 5′-CGATTTTGAGGAAGAACAGC-3′). To modify the targeted locus, we cotransfected a 120-base-pair repair oligonucleotide containing the desired mutation and a silent mutation (repair oligonucleotide:

Article

5'-CCAAAAAGAGCAAAGTGCCTTATACAGAGGAGGGCAAAGATGTAGATGTGTCTGTGCGCGATGCTGAAGAAGAACAGCAGGAGGGTCTGCTATCAGAGGTTAATGCAGAGAAAGTGGTTG-3'). pBabePuro was cotransfected in a 10:1 ratio to the pX330. Transfected clones were selected using 2 µg/µl puromycin for 2 days. Colonies were picked when they were clearly visible, gDNA of clones was isolated and mutations were validated by Sanger sequencing.

To target the C terminus of SCC1, a gRNA (primer: 5'-CCAAGGTTCATATTATATA-3') was cloned into px459 V2.0 (Addgene plasmid no 62988). The SCC1-Halo tag HR template was a gift from J. Rhodes⁵⁵. SCC1-Halo cell lines were generated by cotransfection of pX459 and the SCC1-Halo tag HR vector using FuGENE HD Transfection Reagent. Cells were selected with puromycin (2 µg/ml) for 2 days. Colonies were picked when they were clearly visible and validated using western blot analysis and immunofluorescence.

Antibodies

The following antibodies were used for western blots: SMC1 (A300-055A, Bethyl), CTCF (07-729, Millipore and ab128873, Abcam), HSP90 (F-8, Santa Cruz), SCC1 (05-908, Millipore), tubulin (T5168, Sigma) and H4 (05-858, Millipore). All primary antibodies were used at a 1:1,000 dilution, with the exception of HSP90 and tubulin (1:10,000). Secondary antibodies for western blot analysis were used in a 1:2,000 dilution: goat anti-rabbit-PO and goat anti-mouse-PO (DAKO). For ChIP-seq, we used the following antibodies: SCC1 (ab992, Abcam), CTCF (3418S, Cell Signaling) and IgG (I5006, Sigma-Aldrich).

FRAP

Cells were grown on LabTekII-chambered cover glass (Thermo Scientific Nunc). Two days before imaging, cells were transfected with DNA helicase B fragment fused with near-infrared fluorescent protein (DHB-irFP) using FuGENE HD Transfection Reagent. Before imaging, cells were incubated with 300 nM fluorescent Halo tag ligand JF585 for 30 min. Cells were washed 3 times with normal medium and incubated for 1 h to allow exit of excess of ligand. Medium was replaced twice more with prewarmed Leibovitz L-15 medium (Invitrogen). Live-cell imaging was performed on a Leica SP5 confocal microscope with a 63×1.2 NA water objective using the LAS-AF FRAP-Wizard. Before bleaching, five images were taken. Half of the nucleus of G1 cells was photobleached using 6 pulses of 100% transmission of a 561-nm laser. Subsequently, 600 frames were taken every 2 s. Fluorescence intensity was measured in the bleached and unbleached area by user-defined regions using ImageJ v.1.52q, and adjusted by hand for nucleus movement. Measurements were corrected for photobleaching by monitoring a nonbleached cell. Recovery was quantified by calculating the difference in intensity in the bleached and unbleached regions after background correction. Nondiffusive SCC1-Halo (Extended Data Fig. 3f) was quantified by the relative loss in fluorescence intensity in the unbleached region between the first frame postbleaching and five frames prebleaching.

Colony-formation assay

Cells were seeded at equal density and transfected with short interfering (si)RNAs targeting either no oligonucleotide, luciferase, *CTCF* or *SMC1A*. All siRNAs were ON-TARGETplus SMARTpools manufactured by Dharmacon. Transfection was repeated after 3 days, and after an additional 4 days samples were fixed for 10 min with 96% methanol and stained with 0.25% crystal violet. Cells treated by the same protocol were collected for western blot analysis; samples were collected two days before fixation to have enough cells for western blot analysis.

Chromatin fractionation

For the chromatin fractionation experiment shown in Extended Data Fig. 3e, 50 million cells per cell line were collected and fractionation was performed using Subcellular Protein Fractionation Kit for Cultured Cells (78840, Thermo Fisher Scientific) according to the manufacturer's

protocol, with minor changes. The pellet was washed twice after centrifugation. Western blots were performed as previously described¹⁰.

Hi-C

Samples for Hi-C were prepared as previously described¹⁰. Raw sequence data were mapped and processed using HiC-Pro v.2.9⁵⁶ with hg19 as reference. Statistics on the number of valid pairs and percentage of *cis* contacts are summarized in Extended Data Table 3b, c. Replicates 1 and 2 are highly similar, with a reproducibility >0.98 as assessed by HiCRep v.1.8.0⁵⁷, and were subsequently combined into one Hi-C dataset. The valid pair files generated by HiC-Pro were used to create juicebox ready files using juicebox-pre (juicer tools v.0.7.5)⁵⁸. For visualization, contact matrices were ICE-normalized⁵⁹ and counts were normalized for 100 million contacts per sample.

Loops were then called with HICCUPS v.1.11.09¹⁵ at 5-kb, 10-kb and 25-kb resolution. To visualize the genome-wide effect of the introduced *CTCF* mutations in loops, we performed aggregate peak analysis¹⁵ as implemented in GENOVA v.0.9.8 (<https://github.com/robinweide/GENOVA>), using loops that had previously been defined in wild-type HAP1 cells¹⁰. In brief, for a set of loop coordinates a square submatrix is selected such that it is centred on the corresponding coordinates, with a 100-kb flanking region upstream and downstream. These submatrices are then averaged to obtain a mean contact map for these locations.

Similar to the aggregated peak analysis, aggregate TAD analysis was done to visualize how TAD structures are affected by the *CTCF* mutations. For this analysis, we used TADs that had previously been defined for wild-type HAP1 cells¹⁰. In brief, these TADs were called using HiCseg⁶⁰ on 10-kb matrices as input, Poisson distribution, the extended diagonal model and a maximum number of change points of 50. To compensate for TADs of different sizes, the selected regions are resized before averaging the contact maps. These regions are comprised of the TAD itself and a flanking region of half its size. We calculated the insulation score as previously described⁶¹. The insulation score was computed using the implementation of GENOVA, with a rolling window size of 25 kb. The insulation score was then aligned to TAD borders to create heat maps.

For Extended Data Fig. 6, the compartment score was calculated as previously described⁶². In brief, the compartment score is computed per chromosome arm by obtaining the first eigenvector of the observed over expected matrix, minus 1. Then, this eigenvector is multiplied by the square root of its eigenvalue to obtain the compartment score. To correctly orient the scores so that positive values correspond to compartment-A regions, we used the correlation of the compartment score to H3K4me1 peaks in wild-type cells (J.H.I.H. et al., manuscript in preparation).

For Extended Data Fig. 6e, we compared the effect of *CTCF*^{Y226A/F228A} mutation on genome organization to that of CTCF and cohesin degradation¹². Raw Hi-C data from Gene Expression Omnibus (GEO) accession GSE102884 were converted to HiC-Pro format and ICE-normalized. Relative contact probability profiles were generated using GENOVA.

ChIP-seq

Samples for ChIP-seq were prepared and sequenced as previously described¹⁰, with minor changes. The DNA was sheared using Biorupter Pico (Diagenode), 5 cycles of 15-s on and 90-s off. Reads were first trimmed using TrimGalore v.0.6.0⁶³, then mapped to hg19 using Bowtie2 v.2.3.4⁶⁴ with default settings. Bigwig files were generated with DeepTools v.3.1.3⁶⁵ with the following settings: minimum mapping quality of 15, bin length of 10 bp, extending reads to 200 bp and reads per kilobase per million reads normalization.

Peaks were called for all samples using MACS2 v.2.1.1⁶⁶ with default options. Overlaps between the sets of identified peaks across samples were obtained using BEDtools v.2.25.0⁶⁷. Heat maps were generated using DeepTools⁶⁵ for the different sets of peaks identified in the wild-type cell line, excluding those overlapping blacklisted regions of the genome⁶⁸.

CTCF sites shown in Hi-C contact matrices were obtained from a previous publication¹⁰. In brief, these sites were generated by intersecting CTCF peaks with CTCF motifs from JASPAR CORE 2014⁶⁹, using FIMO⁷⁰ to annotate their motif orientation.

ChIP-qPCR

ChIP-qPCR analysis was performed to assess SCC1 and CTCF abundance at specific genomic loci. SCC1 ChIP was performed three times, and on each ChIP three qPCRs were performed in duplicate. A representative qPCR analysis of each ChIP was used for quantification. For CTCF and IgG, two ChIPs were performed in duplicate. Reactions were performed using SYBR No-Rox Mix 2× (BIOLINE) and run on a LightCycler 480 II (Roche). C_t values were determined for input and ChIP samples, and subsequently the ΔC_t value was converted into a percentage of input. The primers are listed in Extended Data Table 3a.

RNA sequencing

Samples for RNA sequencing were prepared and sequenced as previously described¹⁰. Reads were aligned to hg19 using TopHat v.2.1.1⁷¹ and later counted with HTSeq v.0.11.1⁷² using Gencode v.19 gene-build as reference. Differentially expressed genes were identified with DESeq2 v.1.18.135⁷³, with an adjusted P value threshold of 0.05 and considering only protein-coding genes.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Coordinates are available from the PDB under accession number 6QNX for the SA2-SCC1-CTCF complex. The generated Hi-C, RNA sequencing and ChIP-seq data have been deposited in GEO, accession number GSE126637. Any other relevant data are available from the corresponding authors upon reasonable request.

44. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
45. Bowler, M. W. et al. MASSIF-1: a beamline dedicated to the fully automatic characterization and data collection from crystals of biological macromolecules. *J. Synchrotron Radiat.* **22**, 1540–1547 (2015).
46. Svensson, O., Malbet-Monaco, S., Popov, A., Nurizzo, D. & Bowler, M. W. Fully automatic characterization and data collection from crystals of biological macromolecules. *Acta Crystallogr. D* **71**, 1757–1767 (2015).
47. Svensson, O., Gilks, M., Nurizzo, D. & Bowler, M. W. Multi-position data collection and dynamic beam sizing: recent improvements to the automatic data-collection algorithms on MASSIF-1. *Acta Crystallogr. D* **74**, 433–440 (2018).
48. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D* **66**, 133–144 (2010).
49. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).
50. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
51. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
52. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
53. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).

54. Yin, M. et al. Molecular mechanism of directional CTCF recognition of a diverse range of genomic sites. *Cell Res.* **27**, 1365–1377 (2017).
55. Rhodes, J. D. P. et al. Cohesin can remain associated with chromosomes during DNA replication. *Cell Rep.* **20**, 2749–2755 (2017).
56. Servant, N. et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **16**, 259 (2015).
57. Yang, T. et al. HiCRep: assessing the reproducibility of Hi-C data using a stratum-adjusted correlation coefficient. *Genome Res.* **27**, 1939–1949 (2017).
58. Durand, N. C. et al. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **3**, 95–98 (2016).
59. Imakaev, M. et al. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat. Methods* **9**, 999–1003 (2012).
60. Lévy-Leduc, C., Delattre, M., Mary-Huard, T. & Robin, S. Two-dimensional segmentation for analyzing Hi-C data. *Bioinformatics* **30**, i386–i392 (2014).
61. Crane, E. et al. Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* **523**, 240–244 (2015).
62. Flyamer, I. M. et al. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature* **544**, 110–114 (2017).
63. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10 (2011).
64. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
65. Ramirez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).
66. Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X. S. Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.* **7**, 1728–1740 (2012).
67. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
68. Amemiya, H. M., Kundaje, A., & Boyle, A. P. The ENCODE blacklist: identification of problematic regions of the genome. *Sci Rep.* **9**, 9354 (2019).
69. Mathelier, A. et al. JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **42**, D142–D147 (2014).
70. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017–1018 (2011).
71. Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
72. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
73. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
74. Landgraf, C. et al. Protein interaction networks by proteome peptide scanning. *PLoS Biol.* **2**, e14 (2004).

Acknowledgements This work was funded by EMBL. J.H.I.H., Á.S.C. and B.D.R. were supported by an ERC CoG (772471 ‘CohesinLooping’), M.S.v.R. was supported by the Boehringer Ingelheim Fonds and H.T. and E.d.W. were supported by an ERC StG (637587 ‘HAP-PHEN’). H.T. and E.d.W. are part of the Oncode Institute, which is partly financed by the Dutch Cancer Society. We thank the staff at the ESRF beamline Massif-1; T. Gibson for advice concerning short linear motifs; J. Rhodes and K. Nasmyth for reagents and advice on Halo tagging; R. van der Weide for advice and bioinformatic analyses; and R. Kerkhoven and the NKI Genomics Core Facility for sequencing.

Author contributions Y.L. and K.W.M. initiated the project and proposed the CES motif. Y.L. performed biochemical studies and structural analyses with support from K.W.M. J.H.I.H., R.O., M.S.v.R., L.W. and H.T. performed wet-laboratory cell-based experiments and Á.S.C. performed bioinformatic analyses. K.W.M., E.d.W., B.D.R. and D.P. provided supervision. Y.L., K.W.M., B.D.R. and D.P. were involved in conceptualization, project administration and wrote the original and revised draft with input from all authors.

Competing interests The authors declare no competing interests.

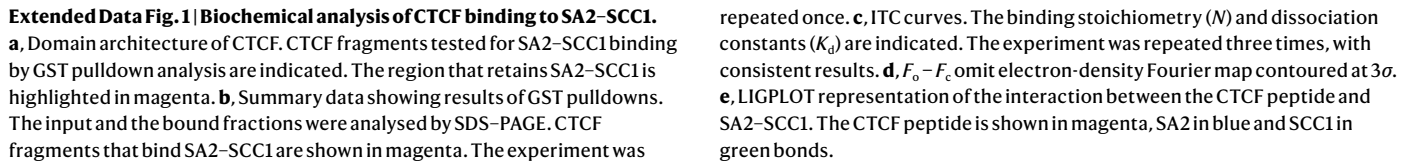
Additional information

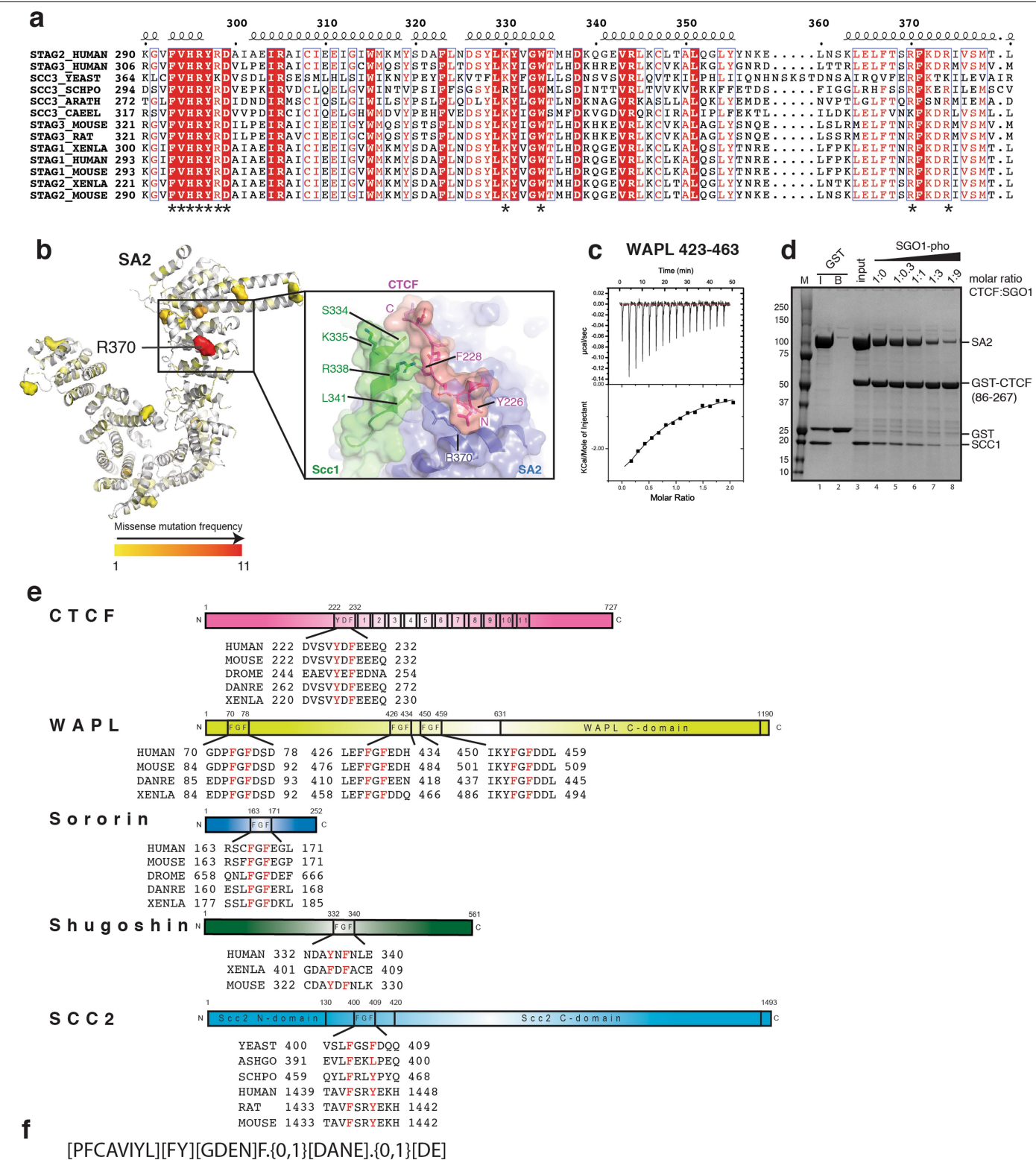
Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-019-1910-z>.

Correspondence and requests for materials should be addressed to K.W.M., E.d.W., B.D.R. or D.P.

Peer review information Nature thanks Victor Corces, Karl-Peter Hopfner and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

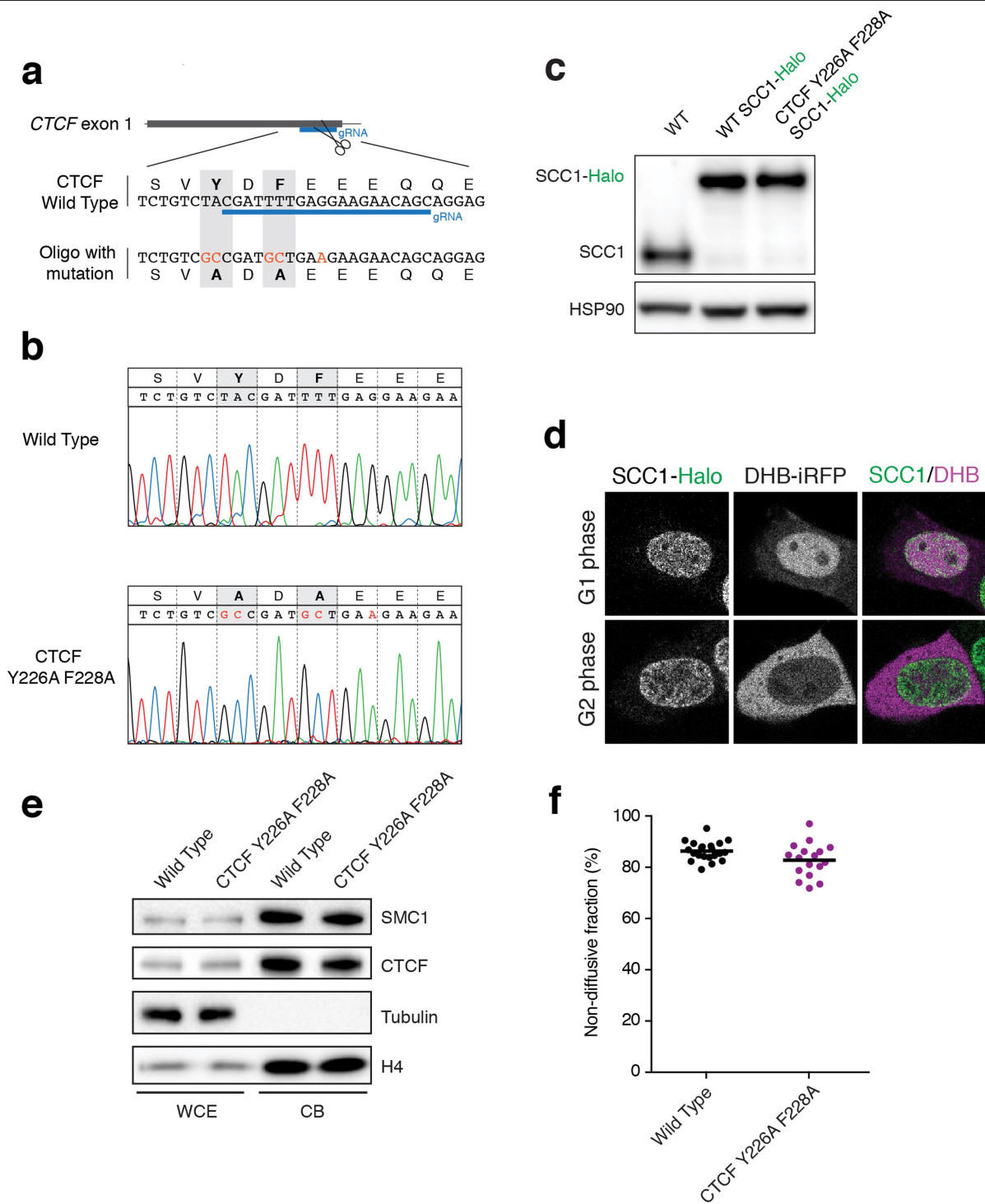
Reprints and permissions information is available at <http://www.nature.com/reprints>.





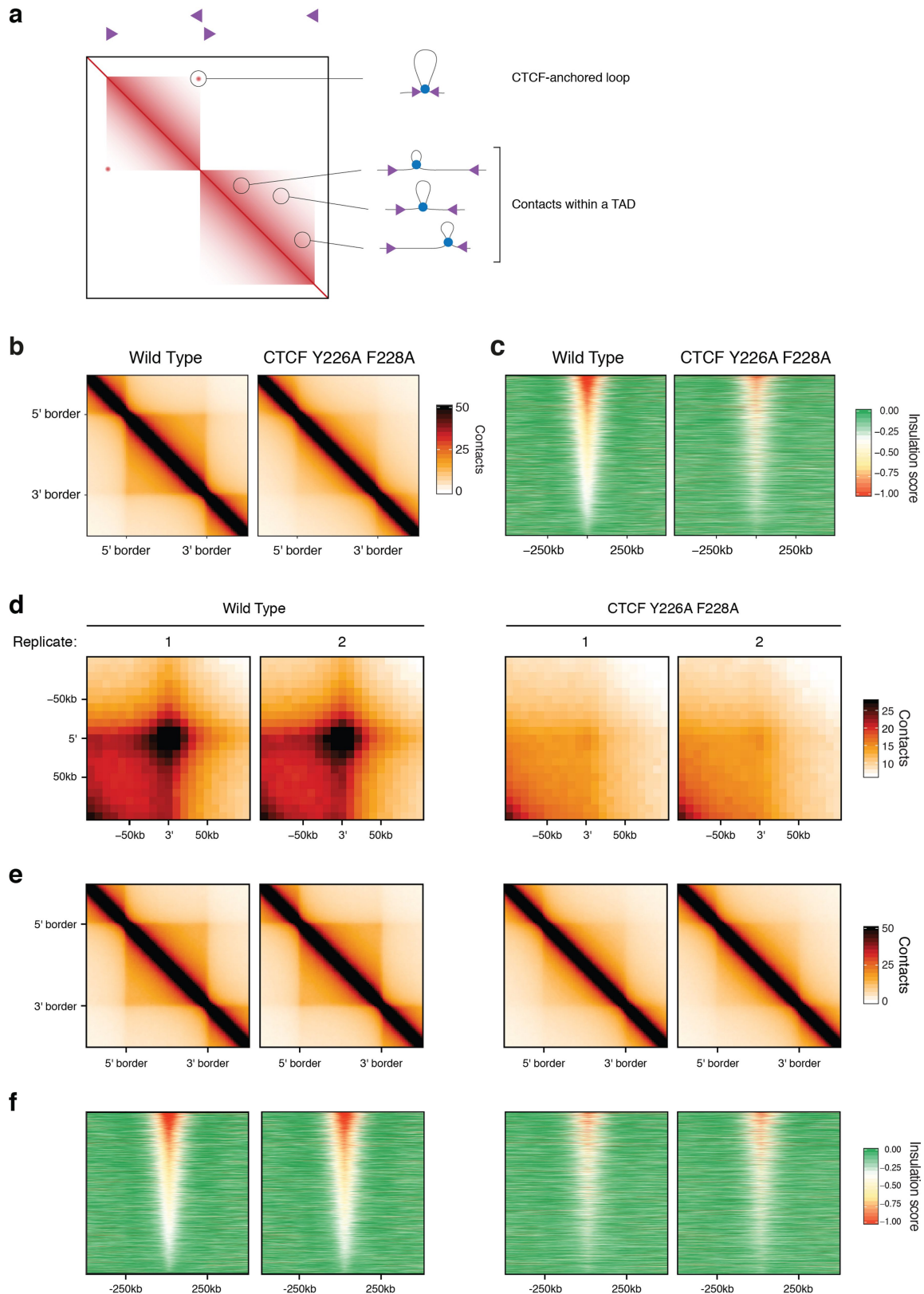
Extended Data Fig. 2 | Analysis of the SA2-SCC1-CTCF structure. a, Multiple sequence alignment of SA2 (here denoted STAG2) or orthologues and paralogues. *Key amino-acid residues that engage CTCF. **b**, Missense mutation frequencies plotted onto the SA2 structure. R370 (a hotspot in SA2) is indicated. The inset shows an overview of the mutation hotspots R370 of SA2, Y226 and F228 of CTCF, and S334, K335, R338 and L341 of SCC1. **c**, ITC progress curves of binding between WAPL(423-463) and SA2-SCC1. **d**, Competition between SGO1 and CTCF for SA2-SCC1 binding. SA2-SCC1 was incubated with GST-CTCF(86-267). Increasing amounts (lanes 4-8) (molar ratios are indicated) of the SGO1

phosphorylated at T346 peptide (spanning residues 331-349) were added and the input and the bound fraction analysed by SDS-PAGE. The experiment was repeated twice. One representative example is shown. **e**, Domain architecture and sequence alignments of cohesin regulators that contain F/YXF motifs. Putative CES-interacting residues are highlighted in red. **f**, Regular expression motif used to query the human and yeast proteomes for factors containing F/YXF motifs. Regular expression syntax: letters denote a specific amino acid; square brackets denote a subset of allowed amino acids; curly brackets denote length variability.



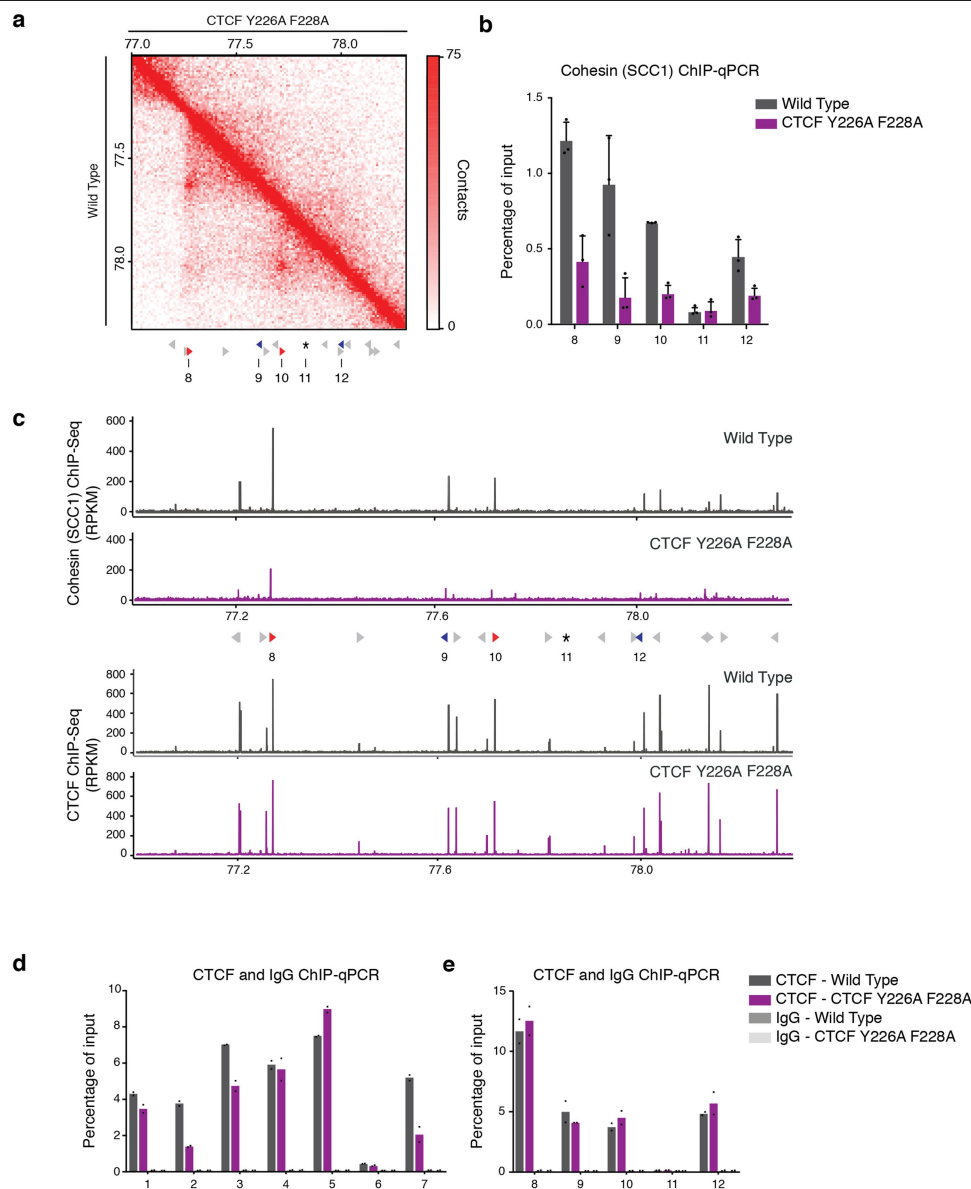
Extended Data Fig. 3 | Generation of *CTCF*^{Y226A/F228A} cells. **a**, Schematic of CRISPR-Cas9-based generation of *CTCF*^{Y226A/F228A} cells. The guide targets cleavage of exon 1 of the *CTCF* gene. The repair oligonucleotide renders the gene noncleavable by Cas9, and simultaneously introduces mutations in the codons that encode Y226 and F228. **b**, The *CTCF*^{Y226A/F228A} mutation was confirmed by Sanger sequencing, including a silent mutation at position 229. **c**, Western blot depicting Halo-tagged SCC1 in wild-type and *CTCF*^{Y226A/F228A} cells. The parental wild-type cells are included as a control. This experiment was performed once. **d**, Representative images of cells in G1 and G2, as indicated by their nuclear and cytoplasmic localization of DHB-iRFP, respectively.

e, Chromatin-bound levels of CTCF and SMC1 analysed by western blot. Histone H4 is used as a control for the chromatin fraction. The *CTCF*^{Y226A/F228A} mutation does not evidently affect overall CTCF and cohesin levels on chromatin. WCE, whole-cell extract; CB, chromatin-bound fraction. This experiment was performed twice with similar results. **f**, Relative SCC1-Halo fluorescence intensity quantified in the unbleached area directly after photobleaching, as a proxy for the chromatin-bound fraction of SCC1. This nondiffusive fraction is not evidently affected by the *CTCF*^{Y226A/F228A} mutation. Individual cells of three independent experiments are plotted as dots and their mean is indicated (21 wild-type cells and 17 *CTCF*^{Y226A/F228A} cells were scored).



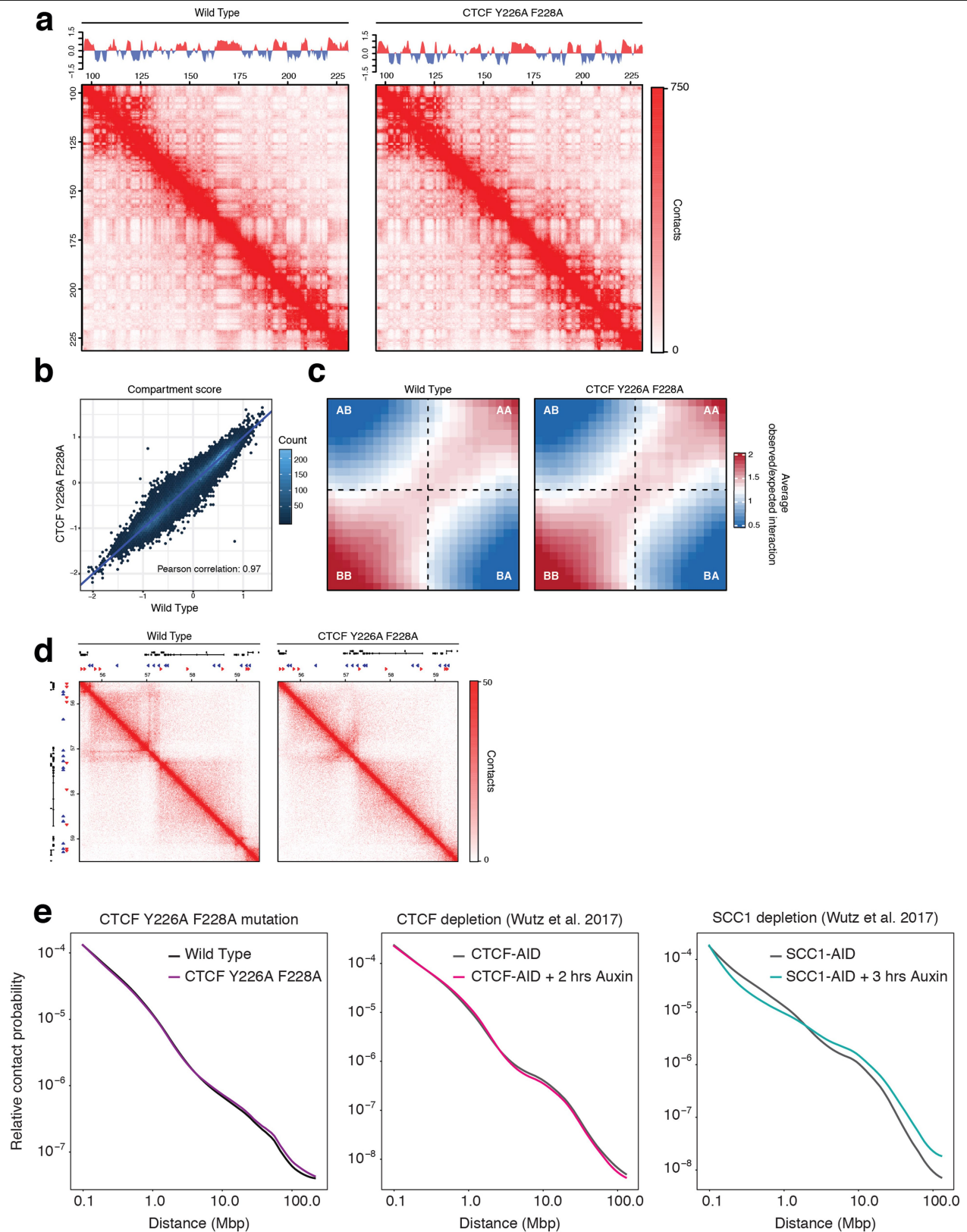
Extended Data Fig. 4 | TAD analyses and Hi-C replicates. **a**, Schematic of a Hi-C matrix displaying DNA-DNA contacts across a genomic region that includes two TADs. TADs in general are flanked by inwards-pointing CTCF sites (magenta arrows). Signal close to the diagonal line reflects short-range contacts, and contacts that span longer distances are found further away from the diagonal. The contacts within a TAD are formed by cohesin complexes (blue circles). Cohesin builds loops that it can enlarge until it encounters CTCF. Some TADs are enriched for contacts between the two CTCF sites that lie at their

boundaries. These contacts are referred to as CTCF-anchored loops. **b**, Aggregate TAD analysis depicting the average contact frequency across TADs defined in wild-type cells. **c**, Heat map of the insulation score⁶¹ at TAD borders, as defined for wild-type cells. **d**, Aggregate peak analysis as in Fig. 3c, using two independent library preparations per genotype. **e**, Aggregate TAD analysis for wild-type and *CTCF*^{Y226A/F228A} cells as in **b**. **f**, Heat map of insulation scores at TAD borders for wild-type and *CTCF*^{Y226A/F228A} cells as in **c**.



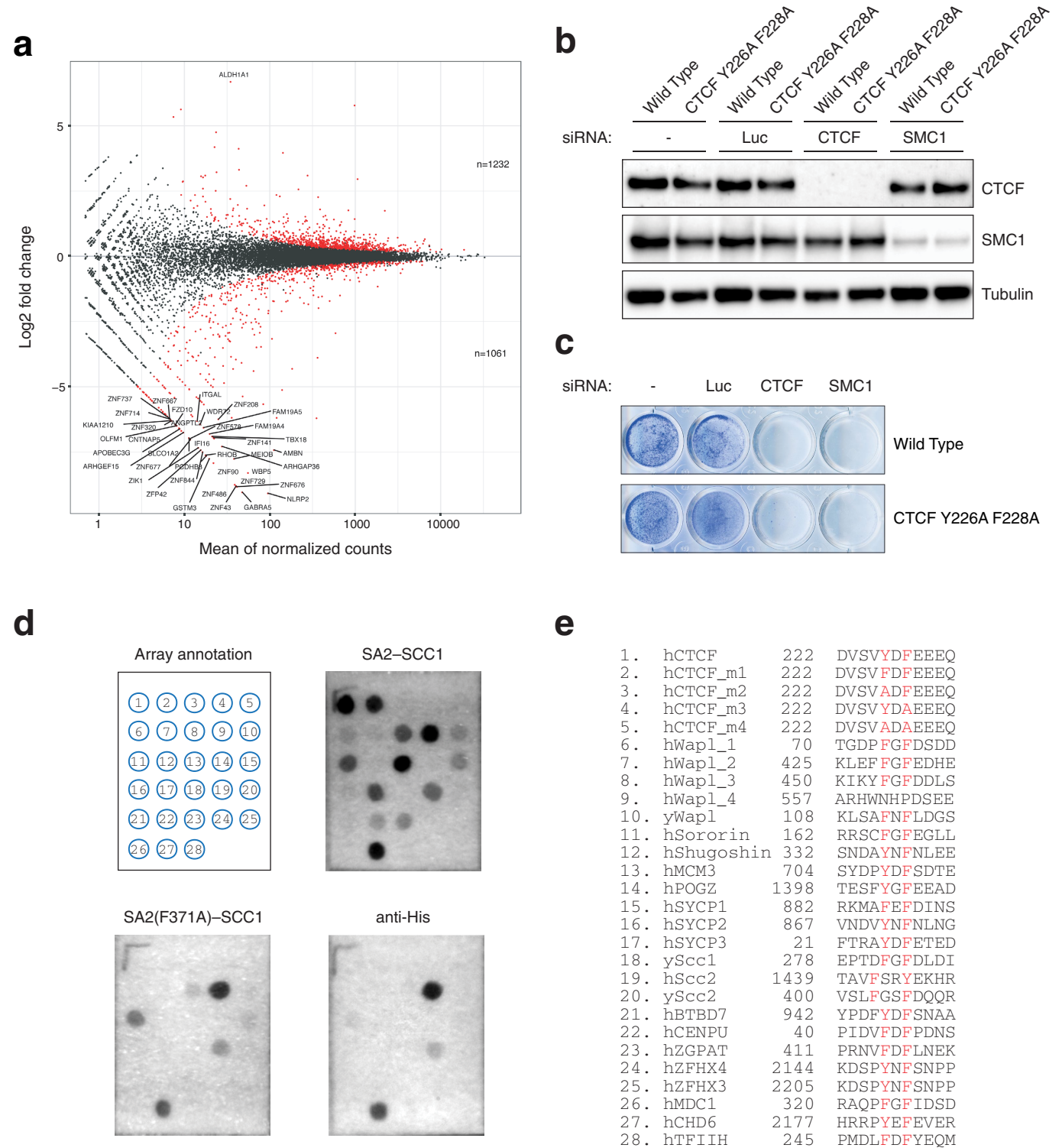
Extended Data Fig. 5 | *CTCF*^{Y226A/F228A} mutation has little effect on CTCF levels at CTCF sites. **a**, Hi-C contact matrix of region chromosome 16: 77000000–78300000 at 10-kb resolution for the wild-type cell line (bottom triangles) and the *CTCF*^{Y226A/F228A} cell line (top triangles). CTCF sites are depicted below; those selected for qPCR are shown in colour. Red triangles indicate sites with a forward motif and blue triangles indicate sites with a reverse motif. The numbers underneath indicate the qPCR primer pairs shown in **b**. Primer pair 11 (indicated with *) is at a locus devoid of SCC1 and CTCF. **b**, ChIP-qPCR analysis of SCC1 (cohesin) enrichment at the aforementioned CTCF sites and control locus (*) in wild-type and *CTCF*^{Y226A/F228A} cells. The mean of three independent

ChIP experiments is shown with the s.d. **c**, ChIP-seq tracks for SCC1 and CTCF at region chromosome 16: 77000000–78300000 in wild-type and *CTCF*^{Y226A/F228A} cells. The loci used for ChIP-qPCR analysis are indicated below the SCC1 ChIP-seq tracks. RPKM, reads per kilobase per million reads. **d**, ChIP-qPCR analysis of CTCF abundance at loci 1–7, as described in Fig. 3d. Analysis includes IgG as a control. The mean of two independent ChIP experiments is shown. Details of replicates are given in the Methods. **e**, ChIP-qPCR analysis of CTCF abundance at loci 8–12, as described in Extended Data Fig. 4a. Analysis includes IgG as a control. The mean of two independent ChIP experiments is shown. Details of replicates are given in the extended methods.



Extended Data Fig. 6 | Compartmentalization is largely maintained in cells that contain the *CTCF*^{Y226A/F228A} mutation. **a**, Hi-C contact matrices of the q-arm of chromosome 2 at 500-kb resolution. The corresponding compartment scores are plotted above. **b**, Genome-wide comparison of compartment scores for wild-type and *CTCF*^{Y226A/F228A} cells. Pearson correlation = 0.97. **c**, Saddle plots representing the interaction between A and B compartments. **d**, A region of

chromosome 1 (55500000–59500000) at 10-kb resolution that contains no obvious CTCF-anchored loops. **e**, Relative contact probability profiles for wild-type and *CTCF*^{Y226A/F228A} mutant cells (left), compared to previously published¹² contact profiles upon degradation of CTCF (middle) or SCC1 (right). The contact probability profile is affected only slightly in the *CTCF*^{Y226A/F228A} mutants, similar to the effects of CTCF depletion.



Extended Data Fig. 7 | Identification of CES ligands. a, Plot depicting the \log_2 -transformed fold change in gene expression in relation to the mean of the normalized counts for each gene. Differentially expressed genes (adjusted P value < 0.05 , two-tailed Wald test adjusted for multiple testing using the Benjamini–Hochberg procedure) are shown in red. Gene names are included for the 40 genes with the highest fold change. **b**, Western blot assessing knockdown of CTCF and the cohesin subunit SMC1 upon transfection with a control siRNA targeting luciferase (luc) or siRNAs targeting *CTCF* or *SMC1A*. This experiment was performed twice with similar results. **c**, Colony formation

assay of wild-type and *CTCF*^{Y226A/F228A} cells upon transfection with a control siRNA targeting luciferase or siRNAs targeting *CTCF* or *SMC1A*. CTCF remains essential for viability in *CTCF*^{Y226A/F228A} cells. This experiment was performed four times with similar results. **d**, Peptide array annotation (top left), binding of SA2–SCC1 (top right) or SA2(F371A)–SCC1 mutant (bottom left) and antibody control (bottom right). Three independent experiments were done, with consistent results. One representative example is shown. **e**, Amino acid sequences of the peptides. Predicted lead-anchoring residues are coloured red.

Extended Data Table 1 | Summary of ITC data, and X-ray data collection and refinement statistics

Protein	Residues	K _d (μM)	ΔH (kcal/mol)	TΔS (kcal/mol)	ΔG (kcal/mol)	N [‡]
CTCF [#]	222-231	1.04 ± 0.20	-11.08 ± 0.70	-2.92 ± 0.82	-8.16 ± 0.09	0.93 ± 0.04
CTCF [†]	86-267	0.62	-13.16	-4.61	-8.54	0.78
Wapl [†]	423-463	32.8	-6.66	-0.54	-6.11	0.62
Wapl [†]	447-462	78.7	-6.81	-1.20	-5.60	1.00
Shugoshin [†]	331-341	13.5	-10.67	-4.02	-6.64	0.83
Shugoshin [†]	331-349 (pT346) [§]	2.32	-20.00	-12.30	-7.69	0.89

**SA2-Scc1-CTCF
PDB 6QNX**

Data collection	
Space group	P2 ₁ 2 ₁ 2 ₁
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	79.02, 107.25, 176.49
Resolution (Å)	45.81–2.70 [^]
<i>R</i> _{sym} or <i>R</i> _{merge}	6.9 (175)*
<i>I</i> / <i>σI</i>	12.0 (0.8)*
<i>CC</i> 1/2	0.99 (0.33)*
Completeness (%)	99.6 (99.7)*
Redundancy	4.4 (4.3)*
Refinement	
Resolution (Å)	45.81–2.70
<i>R</i> _{work} / <i>R</i> _{free}	0.25 / 0.27
No. reflections	46759
No. atoms	16487
SA2	15088
Scc1	1235
Ligand	140 _{CTCF}
<i>B</i> -factors (mean; Å ²)	
SA2	133.4
Scc1	111.3
Ligand	143.6 _{CTCF}
R.m.s deviations	
Bond lengths (Å)	0.004
Bond angles (°)	0.53

a, ITC data summary. **b**, X-ray data collection and refinement statistics.

[#]Three independent experiments were performed. The mean values ± s.d. are shown.

[†]Experiment was performed once.

[‡]Binding stoichiometry.

[§]pT346, phosphothreonine.

[^]Data derived from one crystal.

*Values in parentheses are for the highest-resolution shell.

Extended Data Table 2 | Quantification of peptide arrays

Position	Protein	species	Mutation	Uniprot	Sequence	K _d [μM]
1	CTCF	Human	Wild-type	P49711	222 DVSVYDFEEEQ	1.04 ± 0.16*
2	CTCF	Human	Y226F	P49711	222 DVSVFDFEEEQ	1.60 ± 0.03†
3	CTCF	Human	Y226A	P49711	222 DVSVADFEEEQ	n.b.
4	CTCF	Human	F228A	P49711	222 DVSVYDAEEEQ	n.b.
5	CTCF	Human	Y226/F228A	P49711	222 DVSVADAEEEQ	n.b.
6	WAPL	Human	Wild-type	Q7Z5K2	70 TGDFFGFDSDD	5.90 ± 0.95†
7	WAPL	Human	Wild-type	Q7Z5K2	425 KLEFFGFEDHE	12.43 ± 7.72†
8	WAPL	Human	Wild-type	Q7Z5K2	450 KIKYFGFDDLS	2.17 ± 0.16†
9	WAPL	Human	Wild-type	Q7Z5K2	557 ARHWNHPDSEE	ND
10	WAPL	Yeast	Wild-type	Q99359	108 KLSAFNFDLGS	10.65 ± 3.83†
11	CDCA5	Human	Wild-type	Q96FF9	162 RRSCFGFEGLL	ND
12	SGO1	Human	Wild-type	Q5FBB7	332 SNDAYNFNLEE	n.b.
13	MCM3	Human	Wild-type	P25205	704 SYDPYDFSDE	1.02 ± 0.05†
14	POGZ	Human	Wild-type	Q7Z3K3	1398 TESFYGFEEAD	n.b.
15	SYCP1	Human	Wild-type	Q15431	882 RKMAFEFDINS	6.71 ± 3.18†
16	SYCP2	Human	Wild-type	Q9BX26	867 VNDVYNFNLNG	n.b.
17	SYCP3	Human	Wild-type	Q8IZU3	21 FTRAYDFETED	1.96 ± 0.26†
18	SCC1	Yeast	Wild-type	Q12158	278 EPTDFGFDLDI	n.b.
19	SCC2	Human	Wild-type	Q6KC79	1439 TAVFSRYEKHR	ND
20	SCC2	Yeast	Wild-type	Q04002	400 VSLFGSFDQQR	n.b.
21	BTBD7	Human	Wild-type	Q9P203	942 YPDFYDFSNA	n.b.
22	CENPU	Human	Wild-type	Q71F23	40 PIDVFDFPDNS	4.03 ± 0.22†
23	ZGPAT	Human	Wild-type	Q8N5A5	411 PRNVFDFLNEK	4.03 ± 1.19†
24	ZFHX4	Human	Wild-type	Q86UP3	2144 KDSPYNFSNPP	n.b.
25	ZFHX3	Human	Wild-type	Q15911	2205 KDSPYNFSNPP	n.b.
26	MDC1	Human	Wild-type	Q14676	320 RAQPFGFIDSD	n.b.
27	CHD6	Human	Wild-type	Q8TD26	2177 HRRPYEFEVER	ND
28	TFIIH	Human	Wild-type	Q13888	245 PMDLFDFYEQM	n.b.

Peptide spot signal intensities were correlated to the K_d of CTCF wild type, thus yielding a semiquantitative binding assay⁷⁴. Data points are indicated as mean ± s.d. n.b., no apparent binding. ND, not determined owing to nonspecific binding of the anti-6×histidine antibody.

*Value for CTCF wild type, based on ITC measurement shown in Extended Data Table 1a. The mean values ± s.d. are shown.

†Apparent K_d determined on the basis of three independent peptide array experiments.

Extended Data Table 3 | Primers and Hi-C statistics

Primer set	Primer orientation	Sequence
Primerpair 1	Forward	GGCACTACAGGACCACGTTT
	Reverse	CCCAATTGTGTCTGCCTTTT
Primerpair 2	Forward	GTGGTGTGGGGAAGAGTGTT
	Reverse	GTCAGCTAAACGCCAGGTA
Primerpair 3	Forward	CAAGTTTTCACCCGCTTTA
	Reverse	GAGCCCTAACACCACTCCAC
Primerpair 4	Forward	GGCTTGGAAGCTTTGGTCAT
	Reverse	AGATGGCAGCAGCTTTTCAT
Primerpair 5	Forward	TGATTGTGTACAACAGCTGCAA
	Reverse	ATTTTATAGTGCTCGCAGT
Primerpair 6	Forward	CTGAGCCTCCTGAAAAGTT
	Reverse	CTCTTCTTCGCTCCAGCACT
Primerpair 7	Forward	ACTGCAGCCTCAGCTACCTC
	Reverse	TTTATTGGCATTGCCCTCCTC
Primerpair 8	Forward	CAGTCCTTGTGGCTCCTAGC
	Reverse	TCTGGTGTGCCCTGAACATA
Primerpair 9	Forward	CACCTTGTGGACAGTGGTTG
	Reverse	AGCCTGTGAAACAGGGTGAG
Primerpair 10	Forward	TACACGGGTGGCTAAAGGAG
	Reverse	AGCCAGCCAGATGTCAAACCT
Primerpair 11	Forward	CATGCCCAGCCAATTATTTT
	Reverse	CTCTCCTCCACTTCCCATT
Primerpair 12	Forward	CACCTTTCCGACCCAGAAGA
	Reverse	GGCCTGGAGAAGTCAAACCTG

Genotype	Replicate	Total Pairs	Valid Pairs	Cis	Cis%	Cis < 20kb	Cis > 20kb	Cis ratio
Wild type	1	61118122	60166198	47100811	78,28	7085049	40015762	5,65
Wild type	2.1	62631817	61755440	48127333	77,93	7114243	41043090	5,76
Wild type	2.2	190892790	152708260	122528381	80,24	18087008	104441373	5,77
CTCF Y226A F228A	1	63339779	62197640	47164621	75,83	72900092	39874529	5,47
CTCF Y226A F228A	2.1	62326840	61227593	46569997	76,06	7419962	39150035	5,28
CTCF Y226A F228A	2.2	148586127	118816672	93071165	78,33	14814014	78257151	5,28

Genotype	Replicate	Total Pairs	Valid Pairs	Cis	Cis%	Cis < 20kb	Cis > 20kb	Cis ratio
Wild type	1+2.1+2.2	312814428	270823907	217755919	80,40	32286097	185469852	5,74
CTCF Y226A F228A	1+2.1+2.2	272011360	238342505	186805272	78,38	29523837	157281435	5,33

a, Primers. **b**, Hi-C statistics for replicate library preparations. Libraries 1 and 2 are independent preparations; 2.2 is a deeper resequencing of sample 2.1. The independent libraries 1 and 2.1 were used for Extended Data Fig. 4. A merge of replicates 1, 2.1 and 2.2 of the wild-type cells, and a merge of replicates 1, 2.1 and 2.2 of *CTCF*^{Y226A/F228A} mutant cells, was used for Figs. 3, 4a, Extended Data Figs. 5a, 6. **c**, Hi-C statistics after merging replicates of wild-type and *CTCF*^{Y226A/F228A} libraries.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | | |
|-------------------------------------|--|
| n/a | Confirmed |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	MXCuBE2, XDS (20180808), UNICORN 6, LAS-AF FRAP-Wizard 2.7.4.10100
Data analysis	<p>Molecular replacement was done with Phaser (Phenix 1.14-3260). Hi-C sequencing data was processed with HiC-Pro 2.9. Replicates similarity was assessed with HiCRep 1.8.0. Hi-C data analysis was performed with GENOVA 0.9.98 (github.com/deWitLab/GENOVA). HiC-Pro output was converted to juicer files using juicebox-pre (juicer tools 0.7.5). We performed loop calling with HICCUPS 1.11 on juicer files. ChIPseq reads were trimmed with TrimGalore 0.6.0. Mapping of ChIPseq data was performed with bowtie 2.3.4.130 to hg19. We performed peak calling with MACS2 2.1.131. Overlaps between sets of identified peaks across samples were obtained using BEDtools 2.25.0. ChIPseq alignment plots were created with deeptools 3.1.3. RNAseq data was mapped with TopHat 2.1.133 and count-tables were generated with HTSeq 0.11.1 with the stranded=reverse setting using the Gencode v19 gene-build. Differential expression analysis was performed with DESeq2 1.18.135. Fluorescence intensity in FRAP experiments was measured using ImageJ 1.52q. Structure refinement was done with Phenix (1.14-3260), Structure building was done with COOT 0.8.0-3, Structure renderings were done with Pymol (2.2.3), Structure analysis was done with MolProbity (4.3), Gel band quantification was done with imageJ (1.8.0_112), ITC data were analyzed with Origin 7.0</p>

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Coordinates are available from the Protein Data Bank under accession number 6QNX for the SA2-SCC1-CTCF complex. The generated Hi-C, RNA-Seq and ChIP-Seq data has been deposited in GEO (accession number GSE126637). The current status of these entries is HPUB ('Hold for Publication') which indicates that they are released when the article is published. We will instruct the data depositories to make these entries publicly accessible prior to publication.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. We have added comments on the sample sizes in the legends. For Figures showing GST-pulldown analyses (Fig.1, Extended Data Fig. 1), appropriate controls are used to compare binding side-by-side, as is customary. Single replicates are sufficient in this case. For assays measuring dissociation constants (Extended Data Fig.1, Extended Data Table 1 and 3), three independent experiments were performed as required for determination of measurement errors. Hi-C library preps was performed in duplicate. These replicates showed high similarity and were subsequently merged. RNAseq libraries were generated in triplicate. SCC1 ChIP experiments were performed in triplicate, and CTCF ChIPs in duplicate. These were all analysed by ChIPqPCRs, and a representative of each ChIP was analysed by ChIP-Seq. FRAP was performed on 21 wild type cells, and on 17 CTCF Y226A F228A cells.
Data exclusions	No data was excluded in our analyses.
Replication	We have indicated the number of repeat measurements made and consistency of the results obtained in the figure legends. RNAseq experiments were performed in triplicate. Hi-C was performed in duplicate and data was later pooled together. SCC1 ChIPs were performed in triplicate, and CTCF ChIPs in duplicate. FRAP was performed in three independent experiments. All attempts at replication were successful.
Randomization	Randomization is not relevant to this study, as protein and crystal samples are not required to be allocated into experimental groups. No animals or human research participants are involved in this study.
Blinding	Blinding is not relevant to this study, as protein and crystal samples are not required to be allocated into experimental groups in protein structural studies. No animals or human research participants are involved in this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

Methods

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used

His-HRP antibody (Sigma A-7058 lot number 088M4865V) was used for analysis of peptide arrays at a dilution of 1:2000.

The following primary antibodies were used for ChIP:
 SCC1: Abcam ab992 lot.GR3253930-2 and lot.GR3253930-3 5 ug per ChIP
 CTCF: Cell Signaling 3418S 5 ug per ChIP
 IgG: Sigma-Aldrich I5006 5 ug per ChIP
 For Western blot, the following antibodies were used:
 SMC1: Bethyl A300-055A 6 dilution 1:1000
 SCC1: Millipore 05-908 lot.3055582 dilution 1:1000
 CTCF: Millipore 07-729 lot.3059608 and Abcam ab128873 dilution 1:1000
 HSP90: Santa Cruz F-8 #I0518 dilution 1:10.000
 H4: Millipore 05-858 dilution 1:1000
 Tubulin: Sigma T5168 lot.047M4760V dilution 1:10.000
 Goat anti-Rabbit: DAKO P0447 lot.20046248 dilution 1:2000
 Goat anti-mouse: DAKO P0448 lot.20053537 dilution 1:2000

Validation

<https://www.abcam.com/rad21-antibody-chip-grade-ab992.html>
<https://www.cellsignal.com/products/primary-antibodies/ctcf-d31h2-xp-rabbit-mab/3418>
<https://www.sigmaaldrich.com/catalog/product/sigma/i5006>
<https://www.bethyl.com/product/A300-055A/SMC1+Antibody>
https://www.emdmillipore.com/INTL/en/product/Anti-RAD21-Antibody,MM_NF-05-908
http://www.merckmillipore.com/INTL/en/product/Anti-CTCF-Antibody,MM_NF-07-729
<https://www.abcam.com/ctcf-antibody-epr7314b-ab128873.html>
<https://www.scbt.com/scbt/product/rapgef6-antibody-f-8>
https://www.merckmillipore.com/INTL/en/product/Anti-Histone-H4-Antibody-pan-clone-62-141-13-rabbit-monoclonal,MM_NF-05-858
<https://www.sigmaaldrich.com/catalog/product/sigma/t5168>
https://www.agilent.com/store/en_US/Prod-P044701-2/P044701-2
https://www.agilent.com/store/en_US/Prod-P044801-2/P044801-2

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

HAP1 wild type cells from Carette et al., Nature 2011 a gift from the authors.
 HAP1 CTCF Y226A F228A generated in this study in HAP1 wild type background cells using CRISPR/Cas gene editing.

Authentication

Karyotyping. Mutants were confirmed by Sanger sequencing.

Mycoplasma contamination

All cell lines were negative for mycoplasma contamination.

Commonly misidentified lines (See [ICLAC](#) register)

No commonly misidentified line was used.

ChIP-seq

Data deposition

- ☒ Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links

May remain private before publication.

Go to <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE126637>
 Enter token qzshmsyavvcndmj into the box

Files in database submission

GSM4052950_SCC1_ChIPseq_WT_L001.fastq.gz
 GSM4052950_SCC1_ChIPseq_WT_L002.fastq.gz
 GSM4052950_SCC1_ChIPseq_WT_peaks.narrowPeak.gz
 GSM4052951_SCC1_ChIPseq_CTCF.Y226A.F228A_L001.fastq.gz
 GSM4052951_SCC1_ChIPseq_CTCF.Y226A.F228A_L002.fastq.gz
 GSM4052951_SCC1_ChIPseq_CTCF.Y226A.F228A_peaks.narrowPeak.gz
 GSM4052952_CTCF_ChIPseq_WT_L001.fastq.gz
 GSM4052952_CTCF_ChIPseq_WT_peaks.narrowPeak.gz
 GSM4052953_CTCF_ChIPseq_CTCF.Y226A.F228A_L001.fastq.gz
 GSM4052953_CTCF_ChIPseq_CTCF.Y226A.F228A_peaks.narrowPeak.gz

Genome browser session (e.g. [UCSC](#))

https://genome.ucsc.edu/s/asedeno/CTCF_Y226A_F228A_HAP1

Methodology

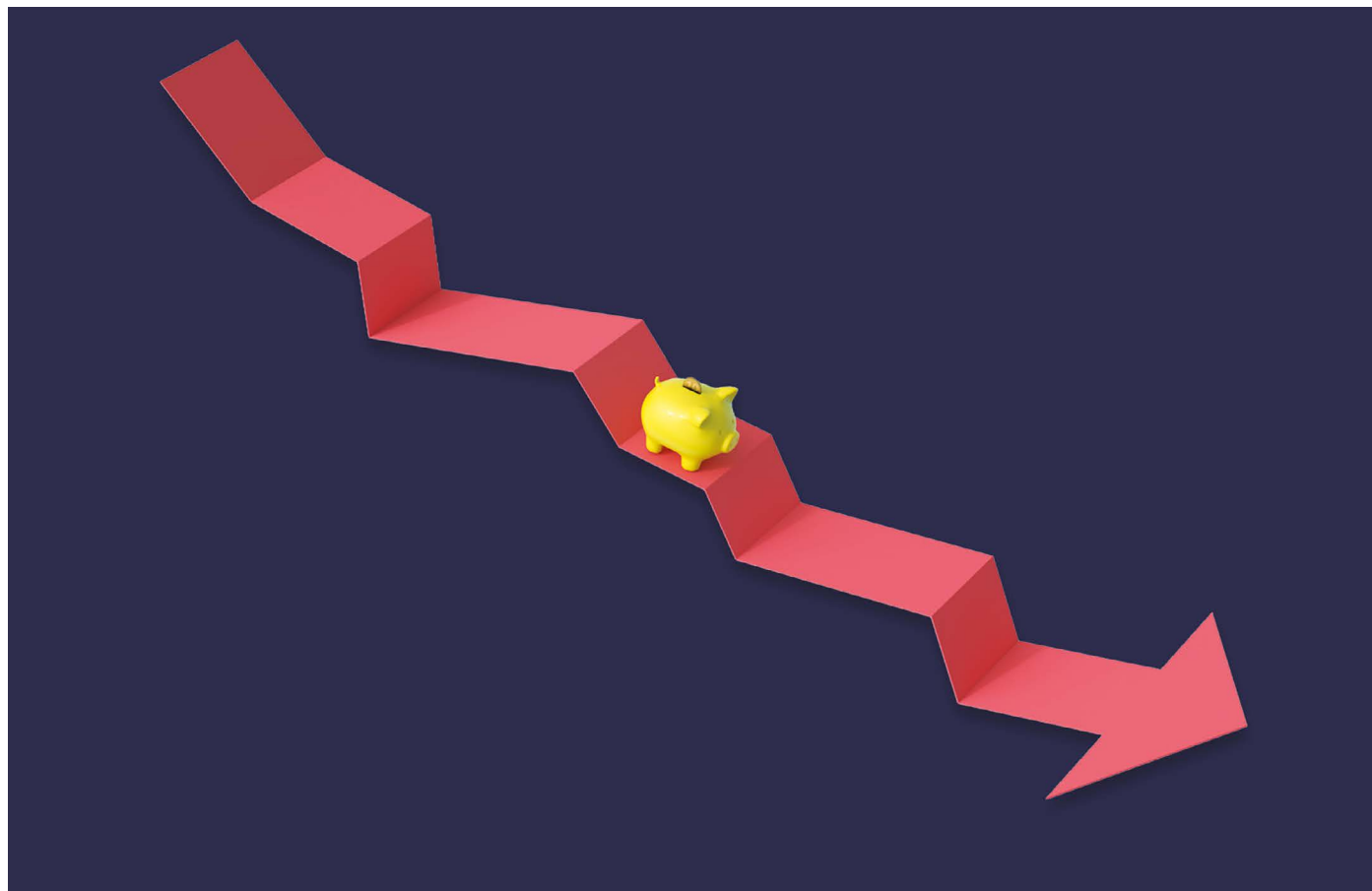
Replicates

SCC1 ChIP experiments were performed in triplicate, and CTCF ChIPs in duplicate. These were all analysed by ChIP-qPCRs, and a representative of each ChIP was analysed by ChIP-Seq

Sequencing depth

sample total_reads uniquely_mapped length type
 5512_11_SCC1_WT_CCGTCC_S25_R1_001 30249392 29761787 65 single

	<div>5512_12_SCC1_CTCFmut_GTGAAA_S26_R1_001 30623586 30175451 65 single 5588_5_CTCF_WT_2_GCCAAT_S146_R1_001 54584946 53634127 65 single 5588_6_CTCF_CTCF103_2_CAGATC_S147_R1_001 54586898 53657421 65 single</div>
Antibodies	<div>SCC1: Abcam ab992 lot.GR3253930-2 and lot.GR3253930-3 CTCF: Cell Signaling 3418S</div>
Peak calling parameters	<div>We performed peak calling with MACS2 2.1.131 for SMC1 and CTCF with standard settings.</div>
Data quality	<div>sample >-log10(0.05) >5FC 5512_11_SCC1_WT_CCGTCC_S25_R1_001_peaks.narrowPeak 52350 29930 5512_12_SCC1_CTCFmut_GTGAAA_S26_R1_001_peaks.narrowPeak 47710 19299 5588_5_CTCF_WT_2_GCCAAT_S146_R1_001_peaks.narrowPeak 71900 57656 5588_6_CTCF_CTCF103_2_CAGATC_S147_R1_001_peaks.narrowPeak 84009 67305</div>
Software	<div>We performed peak calling with MACS2 2.1.131 for SMC1 and CTCF with standard settings.</div>



ADAPTED FROM GETTY

AFTER THE FALL: WHAT TO DO WHEN YOUR GRANT IS REJECTED

Failed funding applications are inevitable, but perseverance can pay dividends. **By James Mitchell Crow**

The day after she submitted a grant proposal last November, Sarah McNaughton listed all the tactics she could think of to boost her chances of success next time. “I expect to be rejected,” says McNaughton. “It is the exception to get funded, not the rule.” Publishing key papers and forging new collaborations were on her list, as was collecting preliminary data.

McNaughton, a nutrition researcher at Deakin University in Melbourne, Australia, studies dietary patterns to find ways to improve public health. For the next phase of her work, she wants volunteers to use wearable cameras to capture what influences their

food choices in real life, so she can determine how those vary depending on a person’s nutrition knowledge and cooking skills.

After McNaughton had sent off her grant application to Australia’s National Health and Medical Research Council (NHMRC), top of her to-do list was launching a pilot study. “If we can show that people will wear the cameras, and they capture the data we need, that would really strengthen the application,” she says.

A good idea is no guarantee of grant success. At the US National Science Foundation (NSF) in 2017 – the most recent year for which data are available – proposals worth

a total of almost US\$4 billion were rejected simply because they were beyond the organization’s budget, even though reviewers had rated them as very good or excellent. At the US National Institutes of Health, the aggregate success rate for research grants was 20.5% in 2017 (the most recent data available). At the biomedical-research funder Wellcome in London, roughly 50% of applications make it through the preliminary stage. Of those, around 20% were funded in 2017–18. And the NHMRC Investigator Grant category that McNaughton applied for had a success rate of just 7% in the previous round in 2019.

“Given the low success rates of funding

More on rejection recovery

It's painful when your grant application is rejected, but here are some further thoughts on helping you to work productively after you've recovered from your disappointment.

- You're not alone. Average success rates are around 20% among large funders, so grant rejection is common. "Don't lose heart," says Shahid Jameel, chief executive of IndiaAlliance, a biomedical-research funder in New Delhi and Hyderabad. Rejection doesn't mean that your work is flawed.
- Give yourself time. Allow a week or so to recover, says Candace Hassall, head of researcher affairs at the biomedical funder Wellcome in London. "When people are turned down, they are angry and upset. Let that play out," she says. Put the application to one side for a few days before you consider your next steps.
- Share your setback. Discussing the grant rejection with colleagues, mentors and others can provide emotional support in the short term, and give you constructive feedback to help you to reapply for the grant when you are ready. "People whose grants have been rejected might not want to tell anybody, but getting advice and input can really help," says Karen Noble, head of research careers at Cancer Research UK, which funds scientists and health-care professionals working on cancer treatments.
- Look for ways to improve. Tackling the concerns of the reviewers who rejected your grant is important. "But don't assume that just by addressing the issues outlined, you will necessarily be successful next time," says Noble. It is unlikely that the same reviewers will see your application again, so look at it holistically and strengthen it for the next round. This might involve incorporating key new data, learning a crucial technique or forming a fresh collaboration.
- Get feedback. Your revision needs review by a broad, diverse group of people, including colleagues, mentors and members of your network. You should also circulate the revision to scientists who don't specialize in your field.

around the world, the odds are stacked against you in winning that one proposal," says Drew Evans, an energy researcher at the University of South Australia in Adelaide, and former deputy chair of the Australian Early- and Mid-Career Researcher Forum. "Work towards a portfolio of activities," he says. Aiming for different strands of funding to cover various aspects of a researcher's work is a safer bet than seeking one major grant, he adds.

McNaughton applies the same strategy to any research for which she is seeking funding. "I think about how I can split it up and target it to other organizations," she says. It's the first step towards applying to different funders without having to start from scratch each time – and you can work on it while waiting for the outcome of one application. "Rather than writing eight different grants, you are building an area – calling on the same literature and on your same publications," McNaughton says.

Planning for rejection is a crucial part of the granting process, say those who have been through the wringer (see 'More on rejection recovery'). The limited pot of research funds worldwide means that competition is fierce. "We receive many more proposals – many more very good proposals – than we can possibly fund," says Dawn Tilbury, a mechanical engineer at the University of Michigan in Ann Arbor who is head of the NSF Engineering Directorate, which funds basic research in science and engineering.

Rejection hurts

Rejection can be a bruising experience, say veteran grant-writers, and applicants need to give themselves at least a week to get through the initial pain. "Take a deep breath, close your computer, go home. Talk to your partner, or pet your cat," says Tilbury. It's a rollercoaster that Evans has ridden plenty of times. "You go through the various stages of emotions – anger, disappointment, despair, grieving almost," he says. "Having time to digest, to get upset and angry – you need to go through that process, because you need a clear mind to come back to it constructively."

But grant-seekers can develop tricks to handle rejection better, says McNaughton. "Part of the reason I make a to-do list is to pull back my expectations," she says. "Once it might have taken me a week or two to bounce back. Now, it's 24 hours."

During the emotional recalibration process, researchers should share the setback with others, including colleagues and other professional contacts, says Evans. "It is your network that is going to give you the support and encouragement to keep going," he says. Peers and mentors can help to put the rejection into context. They might also know of other funding opportunities that can help to bridge an immediate financial shortfall, or

of potential collaborators who might be able to bring a researcher into a larger funding opportunity.

Ask the funder

After working through the emotional component, applicants should next seek feedback from the granting organization. The level of feedback sent out with rejection letters varies drastically, depending on the organization, the scheme applied for and the stage the application reached before rejection.

For smaller funders, feedback might not be provided as a matter of course. "That takes a bit of effort to put together," says Kristina Elvidge, research manager at the Sanfilippo Children's Foundation in Australia. The charity, based near Sydney, funds up to Aus\$700,000 (around US\$472,000) annually on research into treatments for the rare genetic disorder Sanfilippo syndrome, which causes fatal brain damage.

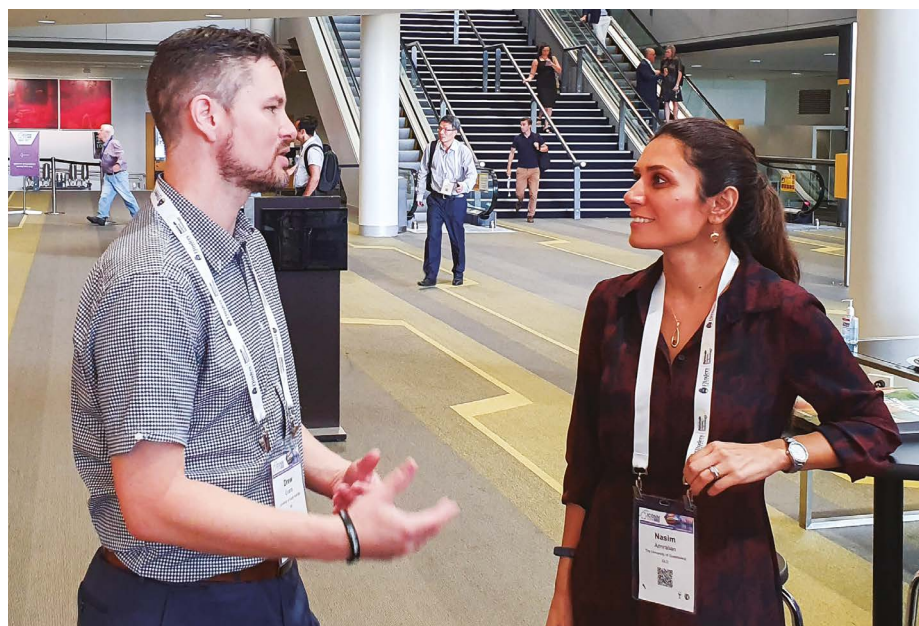
"I always give feedback to rejected applicants if they ask – but they very rarely do," Elvidge says. For researchers whose work might align closely with the mission of a small foundation, seeking feedback can be the first step in starting a dialogue and building a relationship with a potential long-term funder. Megan Donnell, the foundation's executive director and founder, says that the organization welcomes such efforts.

For applicants to a larger organization or agency, such as the NSF, a rejection typically comes with some feedback – but that doesn't mean the researcher can't seek more, Tilbury says. "The programme director might be able to fill in some of the blanks," she says. The feedback can contain many comments, criticisms and suggestions, and often the grant reviewers do not agree with each other. The programme director can help the applicant to peel away superficial concerns and make sure that she or he understands the proposal's underlying weaknesses so as to address them in a potential revision, Tilbury says. "It's one of the things programme directors enjoy doing – mentoring junior faculty members and trying to help them in their grant writing."

Some funders will not have the resources to provide feedback. But researchers should not fear tainting their reputation if they ask, says Candace Hassall, head of researcher affairs at Wellcome. "A funding agency won't think badly of anyone contacting them for advice, even if we can't give it."

Get feedback on the feedback

Once a researcher has gathered constructive criticism, he or she should candidly appraise the strengths and weaknesses of their application. It is important to avoid taking feedback personally, says Shahid Jameel, chief executive of IndiaAlliance, a large research funder in New Delhi and Hyderabad. It supports



Discussing grant rejections with peers can help to put them into context, advises Drew Evans (left), shown talking to early-career researcher Nasim Amiralian.

biomedical and health research in India, and is itself funded by Wellcome and the Indian government's Department of Biotechnology. "You have to get out of this mindset that there is either something wrong with you, or that people are against you," Jameel says. "Reviewers really want you to do well – that is why they are spending their time reviewing your grant and providing feedback."

Reviewer feedback often seems less negative over time, McNaughton says. "I often colour code my reviewers' comments – green for good and red for bad – and then realize that actually, there are a lot of good things in there as well," she says. "These little things can make the process a bit easier." And getting reviewer feedback is certainly preferable to not getting any, she adds. For her most recent rejection, she received only numerical scores on various components of her grant. "Then it is very difficult to know how to improve the application," she says.

Unsuccessful applicants should also seek input from colleagues and others whose opinions they value. "Talk to your peer group and your mentors – they will have been through the process and they can help you interpret the letter," says Karen Noble, head of research careers at Cancer Research UK in London, which funds work on cancer treatments. Researchers can ask colleagues whether they agree with the feedback, whether they think that the reviewers missed an important point because it was not fully explained in the proposal, or whether they consider the proposal's argument to be flawed.

Researchers also need to determine whether they should reapply to the same funding scheme or seek alternatives. If an application fell at the first round of screening – in which

reviewers assess the overall suitability of an applicant and proposal for that particular scheme – an alternative funder could be a better fit. For example, some government-supported agencies, such as the NSF, give grants for only basic research, whereas others, such as the US Department of Energy, are mission-focused and fund more-applied projects. "It is also important to consider funders that are not in one's own nation," says Jameel.

"Industry partnerships are now one of the hot topics around the water cooler."

Grant-writers should keep industrial funders in mind, Evans says. He notes that applicants might be able to reshape a proposal to show its value to a particular business, adding that scientists who engage with businesses can diversify their grant portfolio and boost the resilience of their research income stream. Exploring potential applications of one's work to industry could keep a researcher going until the next round of funding agency grants. "Industry partnerships are now one of the hot topics around the water cooler," he says.

Nailing the details

Rejection also lurks after the preliminary screening stage when a grant application enters peer review. "If there's a particular approach the reviewers don't like, sometimes you may just need to explain it better – but sometimes there's a mismatch," Tilbury says. She adds that many early-career scientists seek to apply a technique or expertise they honed during a postdoc to a new area of research.

If the reviewers weren't sold on the idea, the grant-writer needs to think carefully about the proposal, Tilbury says. "Are the reviewers right? Am I using the wrong hammer to pound this nail?"

If a grant-seeker is certain that their proposal – and their expertise – do fit the grant scheme, they need to make that clear to reviewers. "A common reason for rejection is that the applicant has made assumptions about what the reviewers know about them," Hassall says. "If a technique or method is critical to what you are proposing, you have to include it. Make it easy for people to get the information that they need."

Similarly, if referees rejected a grant because the applicant had no experience in a particular technique, it pays to get it and include that information in the next round. Before reapplying, researchers should seek collaborators who are experts in that area or technique, or spend a week working in the collaborator's lab to gain experience.

It is the applications that just miss out on funding that can be the hardest to revise, Noble says. "Sometimes there wasn't anything inherently wrong with somebody's application. It just didn't make it to the top of the list. Those can be the harder ones to try to repack-age and put in again."

Yet perseverance is key, says Mariane Krause, a psychologist at the Pontifical Catholic University of Chile, and president of the National Commission for Scientific and Technological Research (CONICYT) in Chile, which funds research in the country. She encourages researchers to refine their applications and continue to apply. "I have many young researchers who get a grant the third time," she says.

Reapplying to the same organization for funding can work if the funder allows it. "The success rate of reapplications is significantly higher than for first-time applications," says Alex Martin Hobdley, head of the unit at the European Research Council (ERC) that coordinates project calls and follow-ups. For example, new applicants to ERC grants have a 9–10% success rate. "For people reapplying, the success rate goes up to 14 to 15%. We have people who got their first grant on their seventh application," he adds ([see go.nature.com/2vrfugk](https://go.nature.com/2vrfugk)).

Some schemes impose a specific hiatus period before accepting reapplications, or have an annual or biannual application deadline. Others, including Cancer Research UK, don't impose specific limits. But programme officers recommend resisting the temptation to rush in a revised application as quickly as possible. "Take time – don't knee-jerk," Noble says. "Will you really be in a better position to reapply in a month?"

James Mitchell Crow is a freelance writer in Melbourne, Australia.



Where I work Graham Edgar

As a marine ecologist at the University of Tasmania in Hobart, Australia, I make about 150 dives a year, looking for threatened marine species. I focus on animals and plants that go largely unnoticed: small crustaceans and fish species such as gobies and blennies that grow 3 or 4 centimetres long. I'm trying to illuminate the complex interactions between marine species and to understand how human activities disrupt that process.

This photo was taken in February 2018 at Elizabeth and Middleton Reefs in the southern Coral Sea Islands. I was counting species of fishes and invertebrates, and assessing changes since my last visit in 2013. The silver schooling fish behind me are *Pseudocaranx georgianus*, but I've also seen whale sharks, humpback whales, dolphins and sea snakes. I love the sense of peace underwater: all the stresses and problems of the terrestrial world drop away.

The Reef Life Survey Foundation in Hobart, of which I am president, uses trained volunteer scuba divers to do underwater surveys of biodiversity on rocky and coral

reefs worldwide. The project will provide an irreplaceable record of underwater life on Earth as it is now. We have data from more than 3,000 sites across 53 nations. The work is extremely important because many changes are happening underwater, out of sight. The Great Barrier Reef and Coral Sea reefs are threatened by coral bleaching associated with climate change, cyclones, outbreaks of crown-of-thorns seastars (*Acanthaster planci*) and fishing.

This is the best of jobs in terms of seeing fantastic places that no one has been to before, and knowing that you're recording information needed for dealing with threats to marine biodiversity. But it's depressing to return to places that were once so beautiful to find that climate change, or sediment or pollution, has largely destroyed them. You can't help but reflect on the damage we're doing, and on what a poor state we're leaving Earth in for future generations.

Graham Edgar is a senior marine ecologist at the University of Tasmania in Hobart, Australia.
Interview by Josie Glausiusz.

Photographed by
Scott D. Ling.